DISSERTATION

PLANNING FOR AN UNKNOWN FUTURE: INCORPORATING METEOROLOGICAL UNCERTAINTY

INTO PREDICTIONS OF THE IMPACT OF FIRES AND DUST ON US PARTICULATE MATTER

Submitted by

Steven Brey

Department of Atmospheric Science

Doctoral Committee:

    Advisor: Emily Fischer
    Co-Advisor: Elizabeth Barnes

    Jeffrey Pierce
    Monique Rocca

ABSTRACT

PLANNING FOR AN UNKNOWN FUTURE: INCORPORATING METEOROLOGICAL UNCERTAINTY

INTO PREDICTIONS OF THE IMPACT OF FIRES AND DUST ON US PARTICULATE MATTER

Exposure to particulate matter (PM) pollution has well documented health impacts and is regulated by the United States (U.S.) Environmental Protection Agency (EPA). In the U.S. wildfire smoke and wind-blown dust are significant natural sources of PM pollution. This dissertation shows how the environmental conditions that drive wildfires and wind-blown dust are likely to change in the future and what these changes imply for future PM concentrations.

The first component of this dissertation shows how human ignitions and environmental conditions influence U.S. wildfire activity. Using wildfire burn area and ignition data, I find that in both the western and southeastern U.S., annual lightning- and human-ignited wildfire burn area have similar relationships with key environmental conditions (temperature, relative humidity, and precipitation). These results suggest that burn area for human- and lightning-ignited wildfires will be similarly impacted by climate change. Next, I quantify how the environmental conditions that drive wildfire activity are likely to change in the future under different climate scenarios. Coupled Model Intercomparison Project phase 5 (CMIP5) models agree that western U.S. temperatures will increase in the 21st century for representative concentration pathways (RCPs) 4.5 and 8.5. I find that averaged over seasonal and regional scales, other environmental variables demonstrated to be relevant to fuel flammability and aridity, such as precipitation, evaporation, relative humidity, root zone soil moisture, and wind speed, can be used to explain historical variability in wildfire burn area as well or better than temperature. My work demonstrates that when objectively selecting environmental predictors using Lasso regression, temperature is not always selected, but that this varies by western U.S. ecoregion. When temperature is not selected, the sign and magnitude of future changes in burn area become less certain, highlighting that predicted changes in burn area are sensitive to the environmental predictors chosen to predict burn area. Smaller increases in future wildfire burn area are estimated whenever and wherever the importance of temperature as a predictor is reduced.

The second component of this dissertation examines how environmental conditions that drive fine dust emissions and concentrations in the southwestern U.S. change in the future. I examine environmental conditions that influence dust emissions including, temperature, vapor pressure deficit, relative humidity, precipitation, soil moisture, wind speed, and leaf area index (LAI). My work quantifies fine dust concentrations in the U.S. southwest dust season, March through July, using fine iron as a dust proxy, quantified with measurements from the Interagency Monitoring of PROtected Visual Environments (IMPROVE) network between 1995 and 2015. I show that the largest contribution to the spread in future dust concentration estimates comes from the choice of environmental predictor used to explain observed variability. The spread between different environmental predictor estimates can be larger than the spread between climate scenarios or intermodel spread. Based on linear estimates of how dust concentrations respond to changes in LAI, CMIP5 estimated increases in LAI would result in reduced dust concentrations in the future. However, when I objectively select environmental predictors of dust concentrations using Lasso regression, LAI is not selected in favor of other variables. When using a linear combination of objectively selected environmental variables, I estimate that future southwest dust season mean concentrations will increase by 0.24 μg m^{-3} (12%) by the end of the 21st century for RCP 8.5. This estimated increase in fine dust concentration is driven by decreases in relative humidity, precipitation, soil moisture, and buffered by decreased wind speeds.

ACKNOWLEDGMENTS

I would like to thank Emily Fischer, Libby Barnes, Jeff Pierce, and Monique Rocca for serving on my Masters and PhD committees. Each of you have significantly strengthened this dissertation.

I want to especially thank Emily Fischer for her mentorship over the past five years. It is hard to imagine getting to where I am today without your help. It is also hard to imagine working with a better boss. Thank you.

I want to thank Megan Melamed for her continued mentorship, which started my junior year at the University of Washington.

I want to thank Sarah Tisdale for all of her help making sure I met graduate school and department requirements. Sarah, you are incredible at your job and it makes everyone else's life easier, thank you!

I want to thank three anonymous reviewers for their helpful comments in improving an earlier version of the material in Chapter 3.

I would like to thank Dom Toretto for teaching me to live my life a quarter mile at a time.

Growing up I was never a good student. I always felt like I was moving in slow motion. I failed 9th grade math. I was years behind in reading until my senior year in high school. I scored miserably on the SAT. I got rejected from the state school I wanted to attend three times. I grew up being told that I was not college material. I want to thank those who helped me overcome my academic struggles and gave me a chance to succeed academically. I want to thank John VanLeer at Cascadia Community College for reigniting my passion for weather and showing me that I could succeed. I want to thank David Buchthal. Taking your pre-calculus class at Cascadia rewired my brain in a way that I never thought was possible. I have been successful in every math class I have taken since. I do not think I would have excelled in Math the way I have been able to since 2009 if not for the help of Bryce Goodwin. Bryce, you are a genius, and a good friend, thank you for your help, and showing me how cool math can be. Most importantly, I want to thank Susan Brey (my Mom) for her unconditional support. You always believed that I would succeed if I put in the effort. You also provided me with a secure and loving home, without which, academic success would have eluded me. You provided this security for four children as a widow, demonstrating what real grit is. You are the most important role model in my life. I also want to thank my wife, Kayleigh Wilson. You know better than anyone how much I constantly doubt

myself. You tirelessly battled my negativity with optimism, which helped equip me with the frame of mind required to succeed. Thank you.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION TO WESTERN U.S. PARTICULATE MATTER POLLUTION

Particulate Matter (PM) is a pollutant regulated by the U.S. Environmental Protection Agency (EPA) under the National Ambient Air Quality Standards (NAAQS), a set of requirements mandated by the United States Clean Air Act (Epa and OAR 2014). PM are aerosol that originate from a variety of anthropogenic and natural sources (*e.g.* wind blown dust, wildfire smoke). The composition of PM varies and can depend on its source. Some of the major constituents include sulfate, nitrate, ammonium, organic carbon, sea salt, and dust (Dawson et al. 2014). PM can be emitted directly or formed from gas-phase precursors. When inhaled, $PM_{2.5}$ (particulate matter with diameters less than 2.5 microns) has well documented cardiovascular and respiratory impacts, and exposure can lead to premature death (Correia et al. 2013; Gan et al. 2017; Lipner et al. 2019; Sacks et al. 2011). PM can also influence the earth's radiative budget, heterogeneous chemistry, cloud formation, and visibility (Bond et al. 2013; DeMott et al. 2003; Hand et al. 2016; Krueger et al. 2004). In recent decades, PM concentrations have been declining in many parts of the U.S. These reductions are largely attributed to environmental legislation (Dawson et al. 2014). The public health benefit from these gains in air quality have been significant. Pope et al. (2002) attribute reductions in PM concentrations with nearly 20% of the gains in U.S. life-expectancy between 1980 and 2000. Recently, climate change has put the longevity of recent air quality improvements into question (Dawson et al. 2014). The work in this dissertation is intended to help the scientific community as well as key stakeholders (*e.g.* EPA) better understand the links between climate change and air quality, how they interact, and how they might change in the future. Given the known impacts PM can have on human health and atmospheric composition, understanding how climate change will alter PM air quality is critically important in order to prepare for the future.

Dawson et al. (2014) highlighted research opportunities that would contribute to a better understanding of how PM concentrations and impacts could change in a changing climate. They identified four key areas of research that would help policy makers prepare for the future of PM pollution in the U.S. 1) Improved understanding of the links between climate change and synoptic scale weather phenomena related to PM concentrations (*e.g.* stagnation days, precipitation). 2) A better understanding of the climate drivers of wildfire activity, how wildfire activity impacts PM, and how these might change in the future. 3) Improved estimates of how changing climate and CO_2 concentrations may

affect biogenic volatile organic compounds (VOC) emissions and other PM precursors. 4) A better understanding of the climate drivers of dust emissions and concentrations and how they might change in the future. This dissertation specifically addresses items 2 and 4, towards a better understanding of PM emissions and the air quality impacts that result from wildfire and dust activity. This dissertation focuses on the western U.S, which is home to the largest wildfires and deserts (major dust sources) in the U.S. Although average fine (diameter less than 2.5 microns) PM concentrations are higher in the eastern U.S than the west (van Donkelaar et al. 2012; Van Donkelaar et al. 2010), in the western U.S. extreme fine PM episodes are driven by wildfire smoke (Vedal and Dutton 2006; Ward et al. 2006; Wu et al. 2006), dust events (Claiborn et al. 2000; Lee et al. 2009; Steenburgh et al. 2012), and stagnation (Chen et al. 2012; Malek et al. 2006).

1.2 INTRODUCTION TO WESTERN U.S. WILDFIRES

Since Dawson et al. (2014) outlined a list of key research opportunities that will help prepare the U.S. for an uncertain future, the scientific community has made significant gains in understanding wildfire activity and the ensuing air quality impacts in the western U.S. McClure and Jaffe (2018b) highlighted the continued importance of understanding how wildfire impacts PM air quality; they show that U.S. $PM_{2.5}$ concentrations have only continued to fall outside of areas and seasons not heavily influenced by wildfire smoke. O'Dell et al. (2019) showed that episodic wildfire events impact mean seasonal $PM_{2.5}$ concentrations, which are increasing fire-prone regions (though upward trends are not statistically significant). In the western U.S., the majority of $PM_{2.5}$ NAAQS exceedances can be attributed to wildfire smoke (Liu et al. 2016b). Though the west is home to the largest U.S. wildfires, the entire U.S. airshed is impacted by wildfire smoke. Brey et al. (2018b) showed that western U.S. wildfire smoke is routinely advected thousands of kilometers downwind and that these smoke plumes can elevate ground level $PM_{2.5}$ concentrations. Brey and Fischer (2016) and McClure and Jaffe (2018a) demonstrated that wildfire smoke likely enhances ozone mixing ratios, a gas-phase criteria pollutant.

Wildfires have adverse impacts aside from air quality. Wildfires that overrun suppression efforts, cause damage to property, and result in human fatalities occur every year (Tedim et al. 2018). The 2018 Camp Fire in Butte County California resulted in 86 civilian deaths, the destruction of much of Paradise California, and 16.5 billion USD in total losses, making it the most expensive natural disaster in 2018 (Almukhtar et al. 2018; McBride 2018; Reyes-Velarde 2019). Stand-replacing wildfires (such as the Campfire) can alter carbon balance by turning wildlands into a source for atmospheric carbon

rather than a sink (Kashian et al. 2006). Though this dissertation is focused on understanding how the air quality impacts of wildfires may change in the future, I am also motivated by these other significant impacts.

Progress has been made towards understanding the biophysical and human-controls of wildfire occurrence in the western U.S. Research has shown that western U.S. wildfire activity (*e.g.* burn area) is correlated with temperature, vapor pressure deficit (VPD), hot-dry-windy days, soil moisture, relative humidity, and precipitation (Abatzoglou and Williams 2016; Brey et al. 2018a; Forkel et al. 2017; Kloster and Lasslop 2017; Littell et al. 2009; Parks et al. 2018; Park Williams et al. 2012, 2017; Pechony and Shindell 2010; Rothermel 1983; Short 2015; Srock et al. 2018; Westerling et al. 2006, 2014; Yue et al. 2013a). Modeling studies consistently suggest that wildfire activity (*e.g.* burn area) will continue to increase over the next century Hurteau et al. (2014); Keywood et al. (2013); Moritz et al. (2012); Scholze et al. (2006); Yue et al. (2014). Since the 1800s, humans have had a substantial impact on western U.S. wildfire activity. Wildfire exclusion reduced the amount of fire on the landscape during parts of the past two centuries and has created a fire deficit (Marlon et al. 2012). Humans alter wildfire abundance through exclusion, but also through ignitions. Balch et al. (2017) showed that human-ignited wildfires are expanding when and where wildfires occur.

1.3 INTRODUCTION TO WESTERN U.S. DUST

The sources and impacts of dust PM have also been demonstrated in the scientific literature. Dust is generated from natural (*e.g.* wind over a desert) or anthropogenic (*e.g.* construction or agricultural tilling) abrasive mechanical processes. Dust is a major component of both fine and coarse aerosols in the western U.S. (Li et al. 2013; Wells et al. 2007). U.S. natural dust emissions (wind blown dust of natural non-agriculture lands) are largest in the western U.S., where large deserts and bare ground are located. In the southwest U.S. (an area encompassing Utah, Colorado, Arizona, and New Mexico), local sources of dust include the Mojave, Great Basin, Sonoran, Chihuahuan deserts, and the Southern Great Plains. Dust emissions from these regions are responsible for the highest wind-blown dust concentrations in the U.S. (Prospero 2002). Wind blown desert dust is comprised of mostly silica, which is known to cause chronic lung inflammation, fibrosis, and lung cancer (Prüss-Ustün et al. 2011; Sing and Sing 2010; Steenland and Ward 2014). Dust also impacts visibility in the southwest (Achakulwisut et al. 2018). Hand et al. (2016) showed that spring fine dust concentrations have been increasing in recent years.

Strides have been made in understanding the environmental drivers of dust emissions and the resulting PM concentrations in the western U.S. The presence of vegetation can greatly reduce soil erosion by wind and can limit dust emissions (Kim et al. 2017; Marticorena and Bergametti 1995; Pu and Ginoux 2017; Woodward et al. 2005). Wet (heavy) soil is harder to loft than dry soil, making soil moisture a controlling factor of dust emissions (Fécan et al. 1998). Precipitation can increase soil moisture and scavenge airborne dust (Ginoux et al. 2001). Pu and Ginoux (2017) showed that variations in precipitation, soil bareness, and surface wind speed can control dust emissions in the U.S. great plains. Evan et al. (2016) show that in Australia monthly dust emissions are linearly related to mean monthly wind speed. While observation based analyses reveal key variables that influence dust emissions, process-based models struggle to reproduce observed variability in dust emissions and concentrations (*e.g.* Mahowald et al. (2010); Pu and Ginoux (2018)). Some models in the fifth phase of the Coupled Model Intercomparison Project (CMIP5) are able to capture global mean dust optical depth (Pu and Ginoux 2017). Pu and Ginoux (2017) found that the CMIP5 models that estimate dust emissions tend to estimate a more dusty future than regression analysis, while failing to capture the observed relationships between dust optical depth with surface wind speed, bareness, and precipitation.

1.4   OVERVIEW OF DISSERTATION CHAPTERS AND GOALS

This dissertation documents research towards a better understanding of how the environmental drivers linked to dust and wildfire PM emissions will change in the future and what those changes imply for future emissions and concentrations. My work is a critical step towards preparing for possible future PM concentrations. The overarching strategy implemented in the work that follows is to 1) develop observed relationships between environmental conditions and PM sources and concentrations and, to 2) examine how the environmental conditions that drive PM emissions and concentrations are estimated to change in the future using CMIP5 models. This strategy allows me to estimate possible "future worlds" as estimated by CMIP5 intermodel spread and climate scenarios. This dissertation specifically addresses several key science questions that have emerged since 2014.

Chapter 2 is an article published in *AGU Earth's Future* (Brey et al. 2018a). Chapter 2 provides a comprehensive analysis of how human-ignited wildfires impact when and where wildfires occur, how they impact air quality, how they might change in the future, and contrasts these findings with lightning-ignited wildfires. These questions are explored in western U.S. and southeast U.S. ecoregions, the two

regions responsible for the majority of recent continental U.S. burn area, fuel loading, and emissions (Urbanski et al. 2018). Chapter 2 provides an essential foundation for the work presented in Chapter 3.

Chapter 3 is a manuscript submitted to *AGU Earth's Future*. Chapter 3 uses historical observations of western U.S. wildfire burn area and reanalysis data to identify the environmental conditions that best explain the observed variability in year-to-year summer wildfire burn area in the western U.S. I then identify how these key environmental variables are estimated to change in the future and how these changes are likely to influence future burn area. Estimates of future burn area are quantified across CMIP5 models and two climate forcing scenarios, providing a comprehensive analysis of the spread in possible future summer wildfire burn area, a quantify proportional to but not perfectly correlated to total wildfire emissions.

Chapter 4 is a manuscript in preparation for *AGU Earth's Future*. Chapter 4 uses historical observations of southwest U.S. fine dust concentrations from the Interagency Monitoring of PROtected Visual Environments (IMPROVE) monitoring network as well as reanalysis data to determine what environmental conditions best explain historical year-to-year variability in dust season (March through July) mean fine dust concentrations. I then use CMIP5 model output to estimate how dust emissions may change in the future. These estimates are made for all available CMIP5 models across two climate forcing scenarios. I identify variables that change in a way that would increase future dust concentrations and those that change in ways that would buffer possible increases.

Finally, Chapter 5 summarizes the findings presented throughout the dissertation as well as several possible avenues for future research.

CHAPTER 2

ENVIRONMENTAL CONDITIONS, IGNITION TYPE, AND AIR QUALITY IMPACTS OF WILDFIRES IN THE SOUTHEASTERN AND WESTERN UNITED STATES[1]

2.1 INTRODUCTION

Understanding the current and potential future abundance of fire in the U.S. is important for many reasons, but here we approach this issue from an air quality perspective. Recent work has shown that U.S. wildfires can severely degrade air quality on local to national scales (Baker et al. 2016; Brey et al. 2018b; Brey and Fischer 2016; Ford et al. 2017; Jaffe et al. 2008; Lassman et al. 2017; Saide et al. 2015; Val Martin et al. 2015; Wiedinmyer et al. 2006). For example, on days that exceed $PM_{2.5}$ (mass concentration of particulate matter with diameters smaller than 2.5 μm) regulatory concentrations in the western U.S., 71% of the $PM_{2.5}$ can be attributed to wildfires (Liu et al. 2016b). Brey and Fischer (2016) show that there are many urban areas that experience increased ozone mixing ratios on days impacted by wildfire smoke. Given these air quality impacts, understanding what drives U.S. wildfire abundance and variability is a priority.

Variations in wildfire activity are due to an assortment of influences including, natural climate variability, human-caused climate change, and the legacy of wildfire suppression (Abatzoglou and Williams 2016; Pechony and Shindell 2010; Short 2015; Westerling et al. 2006, 2014). Both long-term and year-to-year climate variability can modulate the availability of fuel in fuel-limited locations (whether or not fuels can survive) and influence fuel flammability in flammability-limited locations (wet fuels do not readily ignite) (Littell et al. 2009, 2016). A large body of literature has demonstrated relationships between western U.S. forests wildfire burn area and meteorology (Abatzoglou and Williams 2016; Abatzoglou and Kolden 2013; Barbero et al. 2014; Cansler and McKenzie 2014; Gannet Hallar et al. 2017; Morton et al. 2013; Park Williams et al. 2015; Riley et al. 2013; Westerling et al. 2014; Yoon et al. 2015). Fewer studies focus on the meteorological drivers of wildfire activity in the southeast U.S., possibly because the majority of burn area in the southeast is from prescribed wildfires (Mitchell et al. 2014) and by design, prescribed wildfire burn area extent is not (normally) determined by meteorology. Approximately 70% of all prescribed burn area in the U.S. occurs in the southeast (Melvin 2015); however, the

---

use of prescribed fire may be limited in the future, as severe droughts in this region become more likely (Mitchell et al. 2014). Unintentional fires occur in this region as well. In a study that focused on Florida wildfires, Brenner (1991) observed that most wildfires between January and April, months were relative humidity values between 20 and 30% are common, are primarily caused by humans, while lightning activity and lightning-ignited wildfire occurrence peaks in July. Given the large number of wildfires in the southeast, their potential importance for regional air quality, and the susceptibility of the region to climate-driven changes in wildfire occurrence (*e.g.* Liu et al. (2013)), this region needs to be understood in order to prepare for the future impacts of U.S. wildfire.

Recently, Short (2014) compiled the Fire Program Analysis Fire Occurrence Data (FPA FOD), one of the most spatially comprehensive U.S. wildfire-occurrence datasets to date. These data include information on the ignition source of wildfires. As a result, there has been an uptick in the number of studies that consider the regionally varying influence and environmental drivers of human-ignited wildfires (*e.g.* Abatzoglou et al. (2016); Balch et al. (2017); Fusco et al. (2016); Nagy et al. (2018); Syphard and Keeley (2015)). The goal of our research is to expand on the findings of these studies by showing how the environmental conditions, meteorological drivers, and air quality impacts differ for human- and lightning-ignited wildfires, and to contrast the southeast and western U.S. (outlined in Figure 2.1). This is an understudied issue essential to understanding what drives wildfire and wildfire-smoke abundance (*i.e.* air quality impacts) in the two continental U.S. regions with the most wildfire activity (based on the FPA FOD, the west and southeast account for 89% of total continental U.S. burn area).

We focus on what drives burn area, rather than number of ignitions, from all wildfires throughout the entire year, including small wildfires. This particular approach is valuable because it reveals several important, but nuanced, new findings. We show that ignoring small wildfires (< 1000 acres), as is often done in similar studies, is not likely appropriate for air-quality-relevant studies. We are able to identify which high-emitting ecoregions are dominated by human-ignited wildfires. Most importantly, we show that annual wildfire burn area for human- and lightning-ignited wildfires within a given ecoregion are correlated with environmental conditions. This implies that changes in burn area that result from climate change will likely be similar for both ignition types.

This paper describes a wildfire occurrence dataset, human- and lightning-ignited wildfire occurrence in the southeast and western US, how these wildfires correlate to each other and environmental conditions, and an estimate of PM2.5 emissions from wildfires of each ignition type in both regions. This work is essential in order for us to understand how the U.S. wildfire system works now and how it

might change in the future. This understanding will help us prepare for wildfire air quality impacts in the future.
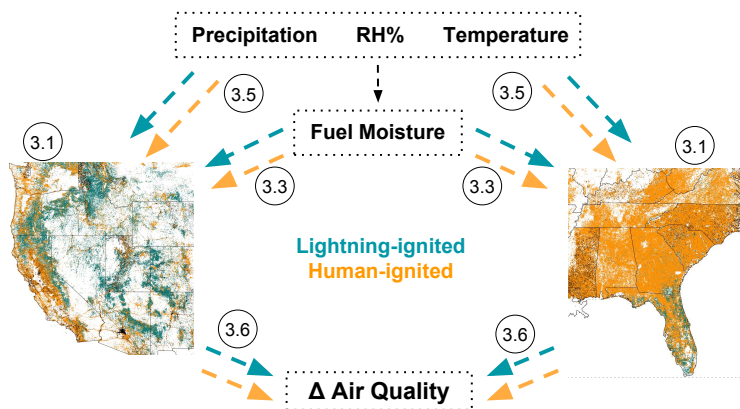

FIG. 2.1. Schematic outlining the components of the U.S. wildfire system explored in this work: wildfire abundance in the west and southeast and how these abundances are influenced by ignition type, precipitation, relative humidity (RH%), temperature, and fuel-moisture. Finally, we show the air quality impacts wildfires have in each region and attribute these impacts to different ignition sources. The numbered circles indicate sections of the paper that correspond to different components of the diagram. Not all sections of the paper are shown. NOTE: The sections labels shown in this schematic match the section numbers of the published manuscript, not this dissertation.

2.2   DATA AND METHODS

2.2.1   Fire Program Analysis Wildfire Occurrence Data (FPA FOD)

The Fire Program Analysis Fire Occurrence Data (hereafter "FPA FOD") documents U.S. wildfires that were extinguished or managed by state, federal, or local agencies between 1992 and 2015 (Short 2014). The database of 1.88 million wildfires merges records from federal, state, and local fire reporting systems. Each wildfire record includes a discovery date, final fire size, cause, and fire location. The wildfire location data are provided as latitude longitude coordinates. However, these locations do not always represent burn scar centroids, ignition locations, nor are they always as precise as the data imply (Short 2014). Sometimes wildfire locations are assigned as a public land survey system (PLSS) section centroid (2.6 km^2) (Short 2014). The specific cause of wildfires is stated, although 8.9% (including Hawaii, Alaska, and Puerto Rico) of the wildfires are assigned a cause of "Missing/Undefined". Other fire causes (*i.e.* ignition sources) in the data include; "Lightning", "Children", "Campfire", "Debris

Burning", "Arson", "Miscellaneous", "Equipment Use", "Smoking", "Railroad", "Fireworks", "Power-line", and "Structure". For this analysis, wildfires with the cause "Missing/Undefined" are discarded or evaluated separately and causes other than "Lightning" are grouped into a single category of "human-ignition". The FPA FOD purposely exclude prescribed (Rx) fires with the exception of prescribed fires that escaped their planned perimeters and turned into wildfires (Short 2014). Fig. A1 in the supplemental information shows the abundance of each ignition type by region and season.

Wildfire characteristics vary significantly between the southeast and western U.S. Figure 2.2 shows the location of Continental U.S. (CONUS) FPA FOD wildfires between 1992 and 2015. In the southeast, wildfires are widespread. There are overlapping wildfires (*e.g.* some lightning-ignited wildfires in the southeast U.S. are hard to see). Figure A2 in the Appendix shows the location of human- and lightning-ignited wildfires on separate maps.
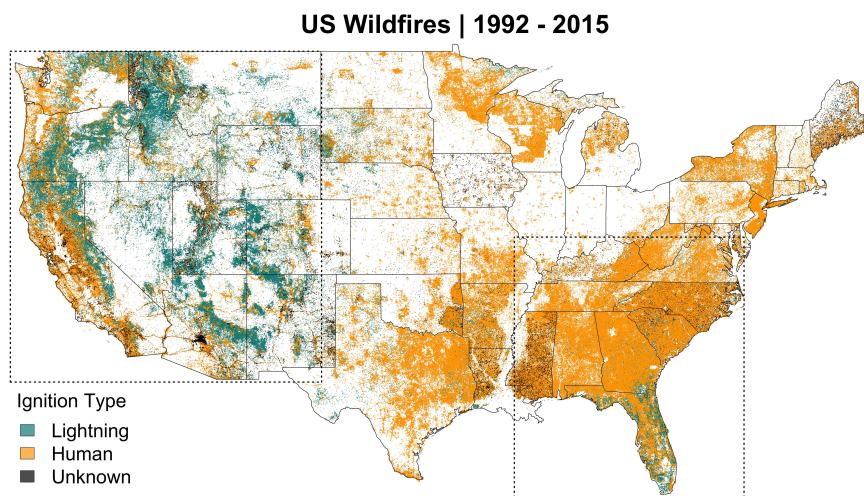

FIG. 2.2. Location and ignition type of all wildfires documented in the Fire Program Analysis Fire Occurrence Data (FPA FOD) located within the contiguous United States between 1992 and 2015. There are 1,880,465 wildfires in the record, 15% ignited by lightning, 77.7% by humans, and 7.3% fires with an Unknown/Missing ignition type. The plotted wildfires account for 107,508,006 acres of area burned over the period. The dashed-line box over the western U.S. shows the geographic extent of the "western U.S." (*i.e.* "west") region. The dashed-line box over the southeast U.S. shows the region defined as "southeast U.S." (*i.e.* "southeast") in this study.

These data are an incomplete record of total wildfire (or any type of fire) activity on the U.S. land-scape (Short 2014). Many wildfires managed or responded to by local agencies go unreported or do not result in a "viable" report (Thomas and Butry 2012). The reliability and quality of these data vary

across states (*e.g.* spatial pattern and abundance of ignition type "Unknown/Undefined", gray dots in Figure 2.2). In some areas it is possible that the ignition type designation "Unknown/Undefined" are wildfires ignited by humans. For example, the Phoenix metropolitan area appears gray in Figure 2.2. Given the high population density of this urban area it is plausible that these wildfires were started by humans rather than lightning.

Short (2014) estimates how complete the FPA FOD are by state with respect to United States Forest Service (USFS) and National Interagency Coordination Center (NICC) wildfire data. These estimates of completeness are provided on a 10 point scale. The scores for states completely within the western region shown in Figure 2.2 range between 7.4 (Colorado) and 9.9 (Idaho), with a mean score of 9.2. The scores for states in the southeast region range between 6.3 (Tennessee) and 9.9 (Georgia), with a mean score of 8.83 (Short 2014). In the west and southeast, the FPA FOD is considered complete to the extent that it can be used to characterize patterns of wildfire area burned (Short 2014).

For this work, the 1992-2015 fire occurrence data were downloaded on 01-29-2018 as a Microsoft Access Database (.accdb) file where the "fires" field was exported. The data are available at the following URL: https://www.fs.usda.gov/rds/archive/Product/RDS-2013-0009.4/.

2.2.2   *North America Level II Ecoregions*

We group the fires in the FPA FOD by ecoregion. Ecoregions are geographic areas where the type of ecosystems are broadly similar. Characteristics considered in making ecosystem classifications include, geology, soils, vegetation, climate, land use, and hydrology (Omernik 1987, 1995). The ecosystems used to separate wildfires in this work were developed by Omernik (1987). These ecosystems serve as a geographic framework for research of areas with similar ecosystem components. Some of the analysis presented in this work is presented on level II ecoregions, a spatial scale useful for subcontinental analysis, and that aggregates wildfires with similar geographic attributes. In this work, FPA FOD wildfires were assigned an ecoregion based on the ecosystem shapefile perimeter in which each fire location fell within, or closest to. The ecoregion shapefile data are available at the following URL: https://www.epa.gov/eco-research/ecoregions-north-america.

2.2.3   *GRIDMet meteorology and fuel moisture information data*

This work uses the GRIDMet derived variable 1000-hour dead fuel moisture to assess fuel aridity. The 1000-hour dead fuel moisture index is determined by ambient environmental conditions and is critical to determining wildfire potential. GRIDMet is a gridded archive of meteorological data over the contiguous U.S. with 4 km horizontal grid spacing. The data were developed by Abatzoglou (2013)

for use in ecological, agriculture, and hydrological models. The data were validated against several weather station networks, including, RAWS, AgriMet, AgWeatherNet, and USHCN-2. These data, as well as documentation are available at the following URL: http://www.climatologylab.org/gridmet.htm

2.2.4 *ECMWF ERA-Interim Reanalysis Meteorology*

We use ECMWF ERA-Interim reanalysis fields (Dee et al. 2011) of temperature, precipitation, and relative humidity to determine regional meteorology conditions at annual temporal scales and horizontal grid spacing of 0.75° x 0.75°. We compute annual means by taking the averages of 6-hourly (00z, 06z, 12z, and 18z) analysis fields. Estimating total precipitation requires the use of forecast fields. We combine the 00z and 12z 12-hour total precipitation fields to get total daily accumulated precipitation. The data are available at https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim

2.2.5 *Global Fire Emissions Data (GFEDv4s)*

In this work, we compare GFEDv4s estimates of burn area to FPA FOD burn area data (Figure A3 in the Appendix shows burn area total estimates for each side by side). FPA FOD are by design an incomplete record of fire activity (no prescribed fire, agricultural burning, fires on private lands, missing reports, *e t c*.). By comparing FPA FOD to the satellite-based GFEDv4s, we can see where FPA FOD wildfires do not capture the majority of satellite detected fire activity. We also use the GFEDv4s estimate of dry fuel consumed to translate regional burn area into total emissions. GFEDv4s better represents the abundance of small fires than previous versions of GFED (*e.g*. GFEDv4). GFEDv4s burn area is based on remotely sensed observations of burn scars. Emissions are estimated by multiplying fire area by an emission factor (emissions per fuel burned) (Giglio et al. 2013; Randerson et al. 2012). The version of GFEDv4s used for this work is gridded on a 0.25 x 0.25° grid. These data are available at http://www.globalfiredata.org/data.html.

2.2.6 *Fire INventory from NCAR (FINN)*

In this work, $PM_{2.5}$ emissions for different regions and individual FPA FOD wildfires are estimated using the Fire INventory from NCAR (FINN) model framework. FINN produces daily fire emission estimates at roughly 1 km spatial resolution (Wiedinmyer et al. 2006, 2011). FINN emissions are estimated using satellite observations of active fires, the fuel loadings at fire locations, biomass consumed, and emission factors for species burned (Wiedinmyer et al. 2006, 2011). Each of these estimated values have associated uncertainties, when combined this results in a factor of two uncertainty associated with FINN emissions estimates (Wiedinmyer et al. 2011). In this work, FPA FOD wildfire location

and year information are used within the FINN model framework instead of MODIS fire detections. Area burned for each fire was assigned based on the FPA FOD data. Updated emission factors and annual-specific MODIS land cover type (2002-2012) and vegetation continuous fields products (2002-2014) were used here to assign land cover and vegetation characteristics. Wildfires where the land cover data indicated the wildfire occurred over water were reassigned as shrubland. Limitations to this method include the assignment to a single land cover and fuel density for each fire based on the latitude and longitude assigned to each fire (not necessarily burn scar centroid). There are uncertainties associated with these emission estimates and it is not clear whether they likely represent over or under-estimates of total emissions for wildfires documented in the FPA FOD. For example, if a large wildfire location is classified in shrubs, but most of the large wildfire ended up burning in a forested area, the emissions would be dramatically underestimated. The opposite could also occur. The emission estimates associated with FPA FOD wildfires used in this work are available at the following URL: ftp://ftp.acom.ucar.edu/user/christin/FPAFOD/.

2.2.7 National Environmental Satellite Data, and Information System (NESDIS) Hazard Mapping System (HMS) HYSPLIT Points

To estimate possible air quality impacts of FPA FOD wildfires, we examine how often wildfires documented in the FPA FOD are co-located with satellite detected fires used to initiate the National Weather Service (NWS) smoke forecasts. The HMS is an interactive environmental satellite image display system where trained satellite analysts identify the geographic locations of fire and smoke in the U.S. (Brey et al. 2018b; Ruminski et al. 2006). Relying primarily on visible satellite imagery, analysts identify the location of fires that they confirm to be producing smoke. These detections are called "HYSPLIT points" because they are used to initiate NOAA's HYSPLIT (Hybrid Single-Particle Lagrangian Integrated Trajectory) model, which is used to estimate the movement of smoke in the NWS smoke forecast (Rolph et al. 2009; Ruminski et al. 2006). The accuracy of the locations assigned to HYSPLIT points is about 2-3 km. The analysis is performed daily. Like many other satellite-based fire-detection methods, smaller wildfires are harder to detect and are more likely to be undercounted than large fires. Analysts make no attempt to distinguish between different types of fires (wildfires, prescribed fires, agriculture fires, etc.). HYSPLIT points with consistent attributes are available from 2007 onward.

2.3.1 *Human versus lightning-ignited wildfires in the west and southeast*
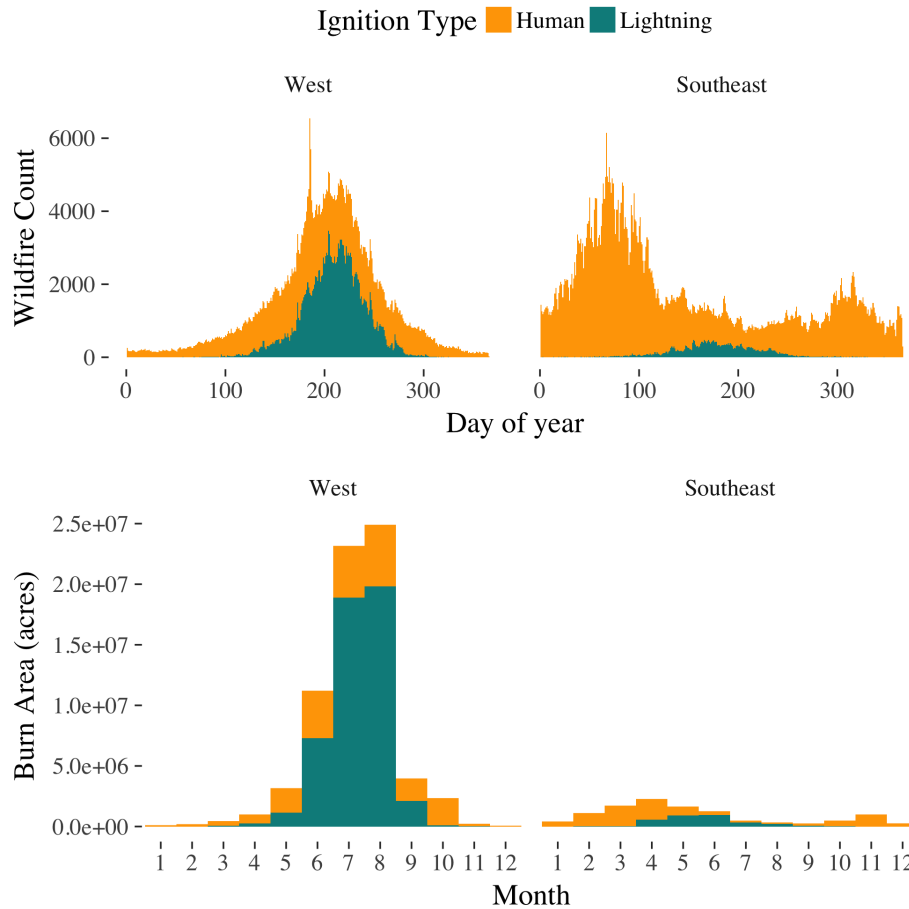

FIG. 2.3. The top row of this figure shows the number of wildfires discovered on a given day of the year in the west (left) and southeast (right) between 1992 and 2015. The bars are color coded by wildfire ignition type. The bottom row shows the monthly burn area that results from the wildfire ignitions shown in the top row. The burn area attributed to a given month is the sum of wildfire burn area for wildfires that started in that month. Wildfires with start causes listed as "Missing/Undefined" have been excluded from this figure. The most common day for ignitions is the 4th of July, an American holiday where the use of fireworks is a common form of celebration.

Previous work has shown that the number of present-day human-ignitions is large, especially in the southeastern U.S. (*e.g.* Balch et al. (2017); Nagy et al. (2018); Syphard et al. (2017b)). We focus on the burn area that results from each ignition type rather than the number of wildfires because burn area is more closely related to the societal impacts of wildfires (air quality, visibility, cost of suppression, carbon cycle, ecological, timber resources, *etc.*). Between 1992 and 2015, the FPA FOD documents 84 million acres of area burned in the west and 12 million in the southeast. In the west, 6% of these acres are

from wildfires with a cause of "Unknown/Undefined" compared to 8% in the southeast. Both regions are impacted by many wildfires, but the direct influence humans have on wildfire abundance is very different between the two regions. In the west, 70% of the area burned is from lightning-ignited wildfires (n=217,187) and 30% is from human-ignited wildfires (n=350,412). In the southeast, 27% of the area burned results from lightning-ignited wildfires (n=37,876) and 73% results from human-ignited wildfires (n=613,013). The large proportion of southeast burn area attributed to human-ignited wildfires suggest that humans are a dominant driver of wildfires activity in this region. Figure A1 shows the regional ignition counts by specific cause and season; differences between the most common types of human ignitions between the southeast and west suggest that there is no single cause that could be jointly eliminated to reduce fires in both regions.

The total number of wildfires during peak wildfire season in the west and southeast are similar (y-axis of top row in Figure 2.3), but there is an order of magnitude more area burned in the west (y-axis of bottom row in Figure 2.3). The southeast accumulates less annual burn area compared to the west, but the burned-area accumulation starts earlier in the year. The southeast wildfire season is bimodal. The first and larger peak occurs early in the year (January, February, March, April), months where parts of the southeast experience low fuel moisture and atmospheric relative humidity (Brenner 1991). These are the months with the largest number of individual wildfires and burn area. These wildfires are almost exclusively ignited by human activities (primarily debris burning). The months May through July account for the majority of lightning-ignited wildfire ignitions and burn area in the southeast. July can be a wet month in much of the southeast. Chiodi et al. (2018) show that for much of the region the probability of daily precipitation exceeding 0.25 inches ranges between 10% and >35%, this can limit the use of prescribed fire. There is a second but smaller peak in the number of human-ignited wildfires in November.

In the west, nearly the opposite occurs. There is a single peak for the number of wildfires and accumulation of burn area for human and lightning-ignited wildfires centered around July and August. Similar to the southeast, in the west the lightning ignited wildfire burn area, as defined as the days of the year that bound the 10 and 90th percentile of the accumulated burn area, is shorter than the human-ignited wildfire season, though there is almost twice the burn area attributed to lightning started wildfires as there are human started wildfires.

The opposite wildfires seasons between the southeast and west means that wildfires have the potential to impact air quality year-round in the southeast. In the late winter and early spring, local wildfires degrade air quality (explored in more detail later). In the summer, when local wildfire activity is minimal, the southeast is regularly impacted by smoke plumes originating from western U.S. wildfires (Brey et al. 2018b). In the fall, when the smoke plumes that originate in the west go away, local human-ignited wildfire activity picks back up.

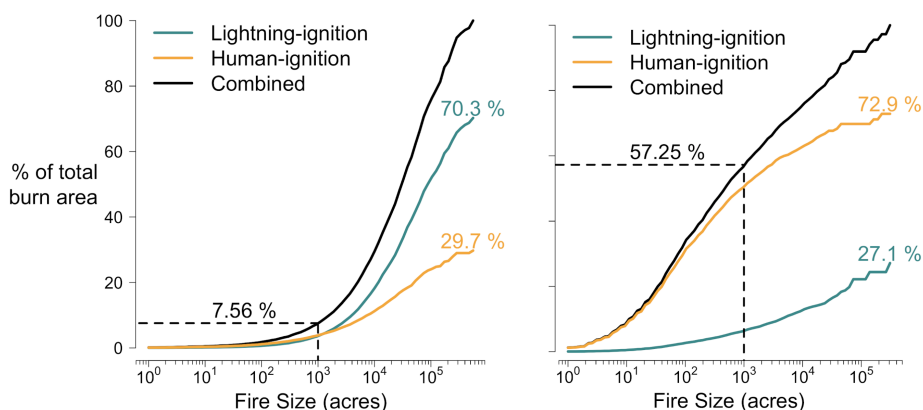### 2.3.2 Ignition type, wildfire size, total burn area, and approximate emissions

FIG. 2.4. Percent of total wildfire burn area (vertical axis) owed to wildfires less than equal to wildfires of a given size (horizontal axis) in the west (left) and southeast (right). The orange and gray curve represents human- and lightning-ignited wildfires respectively. The black curve represents the total of the two. The dashed black line shows the percent of wildfire burn area accounted for by wildfires 1000 acres or smaller (8.54% in the west, 57.25% in the southeast).

The cumulative burn area as a function of wildfire size is presented in Figure 2.4. This demonstrates that large wildfires (> 1000 acres) account for the vast majority of wildfire area burned in the west and account for a smaller proportion in the southeast. This illustrates one of the key differences between wildfires in the southeast and west. The west is home to very large wildfires, and the east is home to smaller wildfires. Wildfires 1000 acres (405 hectares) or smaller account for 8% of the burn area in the west and 57% in the southeast. In the southeast, small human-started wildfires account for half of the wildfire burn area in the region (Figure 2.4, orange line). While in the west, ignoring small wildfires would not significantly change observed patterns of interannual variability, and many studies do this using a similar cutoff (*e.g*. Dennison et al. (2014); Westerling et al. (2014)). The differences in the size of wildfires that contribute to the total burn area for the west and southeast are due to many factors including fuels, climate, frequency of extreme fire weather, topography, population density,

15

road density, and wildfire management priorities. Level II ecoregions provide a scale where some of these factors are broadly similar, thus the following sections show differences between the west and southeast at the scale of level II ecoregions.

Figure 2.5 shows the geographic extent, wildfire burn area, and the dry fuel consumed estimate from GFEDv4s separated by level II ecoregions in the west and southeast. As discussed in the data and methods section, GFEDv4s is a satellite-derived fire emission inventory that includes all satellite detected fires (not just wildfires). This makes comparing GFEDv4s emissions and FPA FOD burn area complicated since there is no assurance that both data sources document the same wildfires. With that said, the GFEDv4s emissions presented in Figure 2.5 offer context for which ecoregions have the highest fire emissions. The forested mountains, the highest emitting ecoregion, consumes an order of magnitude more dry matter than any other ecoregion in the west or southeast (Figure 2.5, bottom row, difference in y-axis).
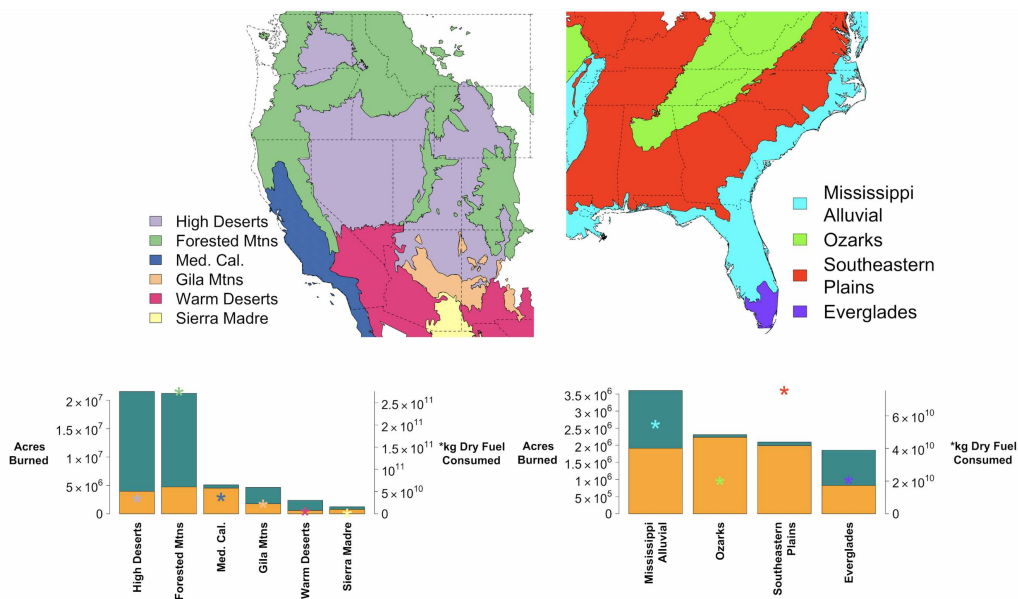

FIG. 2.5. The geographic extent of the ecoregions (top row), total wildfire burn area accounted for by FPA FOD (bars in bottom row, orange and blue represents human-ignited and lightning-ignited burn area respectively), and the all-lands dry fuel consumed by fire estimate from GFEDv4s (asterisk symbol "*" in bottom row) are shown in Figure 5. These totals are shown for level II ecoregions within the west (left) and southeast (right). This figure includes wildfire data between 1997 and 2015 because these are the years FPA FOD and GFEDv4s overlap. The patterns in FPA FOD burn area by region, or percent attributed to human ignitions does not significantly change when including fires from 1992 to 1996. This map shows level II ecoregions. The ecoregions are; forested mountains, high deserts, Mediterranean California, Southeastern Plains, Ozarks, Mississippi alluvial, and Everglades. Areas of white in the maps in the top row are level II ecoregions with very small burn area totals.

In the west, the high desert and forested mountains account for far more burn area than the other seven ecoregions. The fraction of area burned attributed to human-started wildfires in the Forested Mountains and High Deserts is small (orange portion of bars in bottom left of Figure 2.5). Mediterranean California is the only ecoregion in the west where the majority of wildfire burn area are due to human-ignited wildfires. When examining emissions, Figure 2.5 shows that the majority of dry matter consumed in the western U.S. occurs in the Forested Mountains and that most of these emissions are due to lightning-started wildfires.

In the southeast, the Mississippi Alluvial, the Ozarks, the Southeastern Plains, and Everglades account for the majority of the total burn area. The burn area and total emissions within southeast ecoregions is an order of magnitude less than those in the west, though given the abundance of small wildfires in the southeast (*e.g.* Figure 2.4 right panel) and the challenge of detecting small fires with satellites (Randerson et al. 2012), it is possible GFED fire emission estimates in these ecoregions are severely underestimated. Additionally, there is more prescribed fire burn area in the southeast than wildfire burn area (Chiodi et al. 2018; Mitchell et al. 2014), so it is important to keep in mind the large disconnect between total emissions and wildfire burn area presented in the southeast in Figure 2.5.

2.3.3 *Ecoregion-specific relationships between ignition and dead-fuel moisture*

The dead-fuel moisture index is a measure of the amount of water in dead vegetation and is widely used to estimate wildfire potential (NCDC). It is more difficult to ignite fires when fuel moisture levels are high because energy must be used to evaporate water before combustion can occur (Bradshaw et al. 1984; Fosberg et al. 1971; Rothermel 1972; Stocks et al. 1989). Dead-fuel moisture content changes in response to environmental conditions (NCDC), *i.e.* dead-fuel moisture is a quantity with memory of meteorological conditions that drive local moisture abundance (*e.g.* precipitation, temperature, evaporation, *etc.*). The 1000-hour dead-fuel moisture is a measure of the amount of water in fuel as a percentage of its mass where it is assumed that there is a 1000-hour time lag for fuels 3-8 inches in diameter to become $\frac{2}{3}$ of the way towards equilibrium with the local environment. Fuels with small diameters (*e.g.* grass, leaves, twigs) respond to atmospheric moisture levels quickly. FPA FOD wildfires were assigned 1000-hour dead fuel moisture percentages based on the value of the GRIDMet (Abatzoglou 2013) grid cell (4 x 4 km) the wildfire location fell within. For the purpose of assigning fuel moisture values to wildfires, the discovery date is treated as the ignition date. In reality these two dates are not always the same and depending on conditions, can be weeks apart.
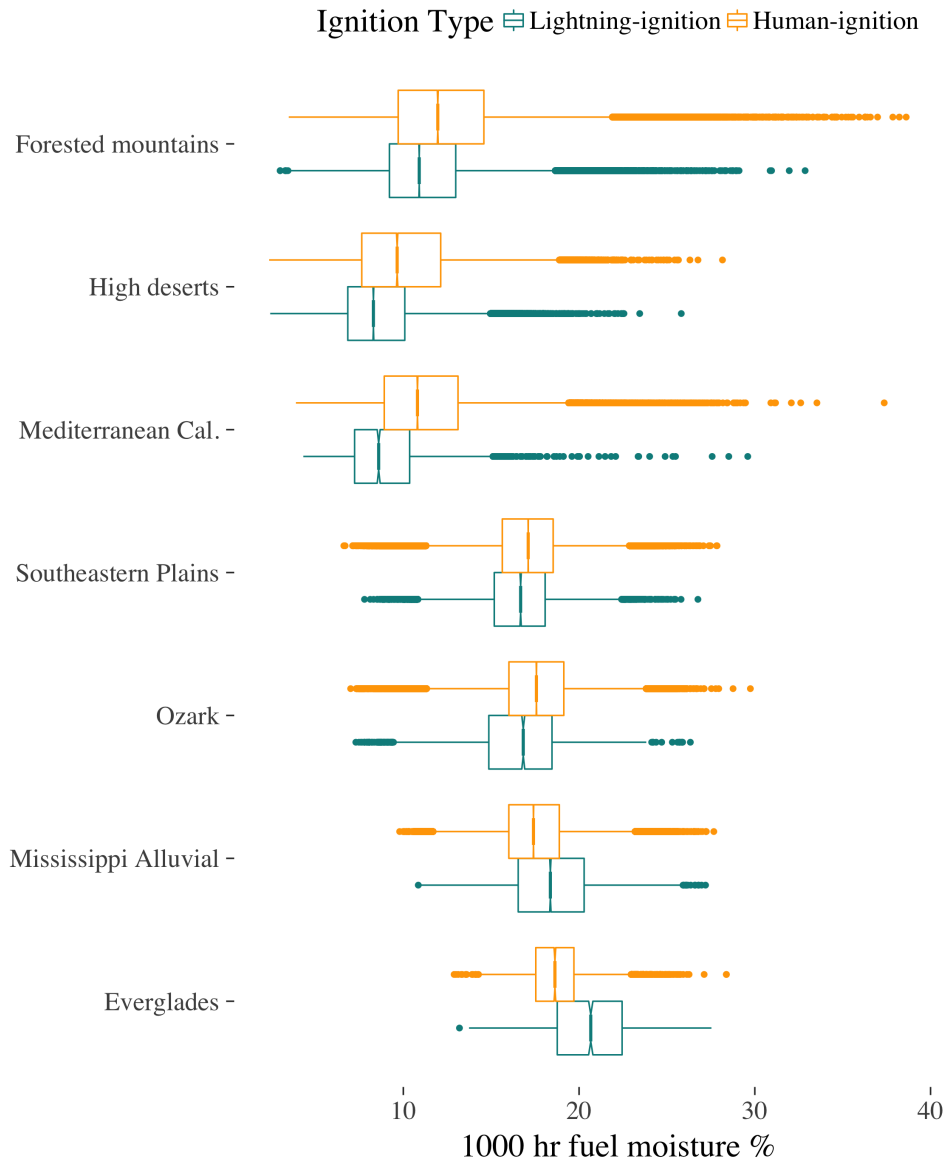
FIG. 2.6. The 1000-hour fuel moisture percent assigned to wildfires based on the location and discovery date in the FPA FOD. Wildfire 1000-hour fuel moisture percent values are separated by level II ecoregions (rows) and ignition type (color). All wildfires (size, month, etc.) were retained for this plot. The boxes show the 25th, 50th (notched middle), and 75th percentiles of the data. The upper whiskers show data that are within 1.5 times the length of the interquartile range (IQR). Data beyond the whiskers are "outlying" points and are plotted individually. The notches around the median value extend 1.58 times the width of the IQR divided by the square root of the number of samples; this gives an approximate way to compare the median values of distributions with roughly 95% confidence (Mcgill et al. 1978).

Figure 2.6 shows the 1000-hour fuel moisture percent assigned to wildfires based on the location and the discovery date provided by the FPA FOD. The box and whisker plots show the distribution of fuel moisture values in the level II ecoregions with the most burn area and emissions for both ignition

types. The most clear distinction in fuel moisture where wildfires occur is the ecoregion, not the type of ignition (rows rather than colors in Figure 2.6). On average, wildfires in the southeast (rows 4 through 7) occur at higher fuel moisture percentages than wildfires in the west (rows 1 - 3). However, due to the abundance of human-ignited wildfires in the southeast and lightning-ignited wildfires in the west, a continental-wide comparison of the fuel moisture percent of wildfires segregated by ignition type would reveal that human-ignited wildfires occur at higher fuel-moisture values than lightning-ignited wildfires. Such a conclusion was made by Balch et al. (2017), who showed that nationally, human-ignited wildfires occur at higher fuel moisture content values than lightning-ignited wildfires. While differences in fuel moisture values between ignition types exist, Figure 2.6 shows that this difference can be largely explained by contrasting ecoregions in the west and southeast.

In the west, the median fuel moisture for wildfires is always higher than for human-ignited wildfires. This is likely because human-ignited wildfires in the west start earlier, go later in the year (Figure 2.3), and are on average small (Figure 2.4). In the southeast ecoregions, the differences between the median fuel-moisture values for each ignition type are generally smaller than the west. In the Everglades and Mississippi Alluvial, the median fuel moisture is higher for lightning-ignited wildfires than human-ignited wildfires. Some of the differences in fuel moisture at wildfire locations in the southeast can be explained by the difference in seasonality for wildfires of each ignition type (bi-modal human-started wildfires season, lightning started wildfire season peaks in the summer). On average, small wildfires (<1000 acres) occur at higher fuel-moisture values and on more days of the year than large wildfires (>1000 acres) (Figure A4). Within a given ecoregion, most large wildfires occur in similar, temperatures, days of the year, and fuel moisture values regardless of ignition type (Figure A4 in the appendix). There can be large interannual variability in wildfire occurrence and the meteorology that drives 1000-hour fuel moisture. This will be explored in the next two sections.

2.3.4   *Interannual variability in lightning- and human-ignited wildfire burn area*

Figure 2.7 shows the Spearman correlation of the interannual variability of lightning and human-ignited burn area for the three ecoregions in the west and four in the southeast that account for the majority of Continental U.S. (CONUS) FPA FOD wildfire activity. We used the Spearman correlation coefficient in Figure 2.7 because we did not want to impose the requirement of a linear relationship between annual human- and lightning-ignited burn area within an ecoregion. For the majority of ecoregions, the interannual variability in burn area between each ignition type are well correlated; however,

correlations are generally higher for ecoregions in the west (light blue in Figure 2.7). The two main exceptions are the Everglades (r = 0.3) and Mediterranean California (r = 0.19). The burn area in Mediterranean California is dominated by human-ignited wildfires while the burn area ignition attribution is almost evenly split in the Everglades, so it is not clear that low correlations are associated with a particular dominant ignition type. Wildfire burn area is dominated by human-ignitions in the Ozarks and Southeastern Plains and the correlation for burn area between ignition types is high in these ecoregions (r > 0.5). The correlation in burn area between ignition types suggests that burn-area totals for both wildfire types are driven by similar factors (*e.g.* environmental conditions) or that the separate factors driving each ignition type are also well correlated.
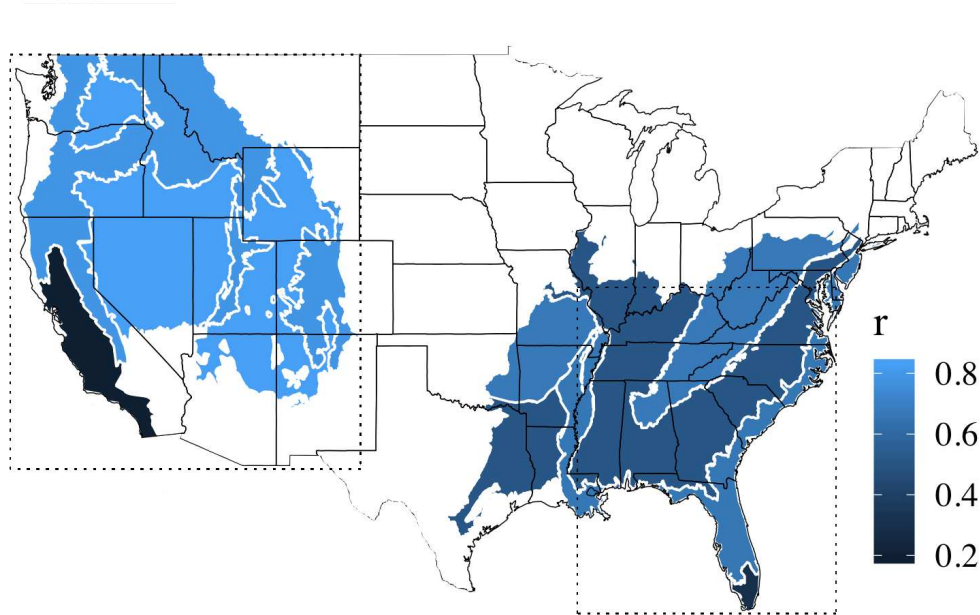

FIG. 2.7. Spearman correlation coefficient (r, mapped as color) for the interannual burn area by ignition type (human and lightning) in each ecoregion between 1992 and 2015. The west and southeast regions are shown with the dashed line boxes, only wildfires that occurred within the plotted ecoregions and boxes were used for the correlation calculations.

2.3.5   *Burn area and annual meteorology across human- and lightning-ignited wildfires*

An important question is whether the variability in ecoregion area burned for lightning- and human-ignited wildfires can be explained by the annual-mean state of meteorological variables. The goal of this section is to determine if the interannual variability of total burn area for human-ignited and lightning-ignited wildfires (shown in Figure 2.7) have the same relationship with key meteorology variables that drive aridity and increase wildfire danger rating indices. Under arid conditions, correlated

indices such as the Palmer drought severity index (PDSI), wildfire energy release component (ERC), and dead-fuel moisture indices all contribute to heightened fire danger (WFAS). Each of these indices are at least partially driven by temperature, precipitation, and relative humidity.
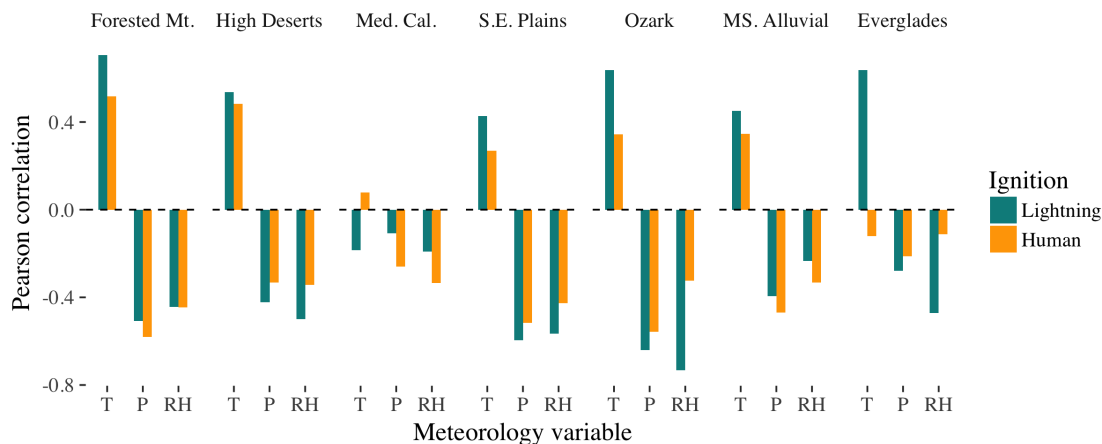

FIG. 2.8. The Pearson (linear) correlation coefficients between total year-long area burned vs. mean annual temperature (T), total precipitation (P), and mean relative humidity percent (RH). Correlation values are shown for burn area by ignition type (color). Temperature and relative humidity are calculated using the mean of daily mean 2-meter temperature for 0.75 x 0.75° ECMWF reanalysis grid boxes that overlap the specified ecoregions. Total precipitation is the annual sum of precipitation for grid boxes that overlap each ecoregion. Wildfires of all sizes were included in the correlation calculations shown in this plot.

Figure 2.8 shows the Pearson correlation coefficient between total summer wildfire burn area and meteorology variables, segregated by ecoregion (columns) and ignition type (color). Relationships (sign of correlation) between burn area and meteorological variables are consistent across ecoregions and ignition types. The exception is temperature in Mediterranean California and the Everglades. In Mediterranean California, temperature is positively correlated with human-ignited burn area and negatively correlated with lightning-ignited burn area. In the Everglades the opposite is true; however, the values of these correlations are close to zero. Correlations are generally stronger for lightning-ignited wildfires, which could suggest that lightning-ignited wildfire burn area is more strongly influenced by meteorology. There is no obvious difference in the magnitudes of the correlations between ecoregions in the west and southeast. This is surprising given the contrasts in wildfire size, fuels, topography, and dominate ignition types between the two regions.

The expectation that total area burned will increase under climate change can be deduced from first principles of fire behavior. Fires ignite more easily and spread faster in dry wildland environments (Albini and Stocks 1986; Alexander and Cruz 2006; Rothermel 1972; Schaaf et al. 2007; Wagner 1977,

1998). All of these ecoregions are located where there is a high degree of confidence that mean surface temperatures will increase over the coming century (Collins et al. 2006). Based on the positive linear correlations between temperature and burn area shown in Figure 2.8 (warmer years tend to have more burn area), it is possible that the future burn area for both ignition types will increase with increasing temperatures (this is not clear in Mediterranean California or the Everglades). However, this predicted increase assumes that, 1) the observed relationships between annual burn area and temperature remain stationary in the future (an assumption Higuera et al. (2015) have shown may not always be true), and 2) these relationships hold when the climate departs from the range of temperatures observed between 1992 and 2015.

Figure 2.8 shows a negative linear relationship between burn area and total annual precipitation in every region for each ignition type. What this anticorrelation implies for changes to wildfire abundance under climate change is limited by the same challenges outlined in discussing temperature, with the added complication that the changes in the quantity and variability of precipitation is not as certain as temperature (Collins et al. 2013; Pendergrass et al. 2017).

This section does not intend to make predictions of future wildfire abundance based on these correlations, but rather to show that whatever climate-driven changes occur in temperature, total precipitation, and relative humidity, Figure 2.8 suggest that the impact on human- and lightning-ignited wildfire burn area may be similar. Syphard et al. (2017a) and Parisien et al. (2016) argue that climate may be less important for driving wildfire activity in areas where human activities and ignitions are pervasive. While humans are responsible for igniting the majority of wildfires in the southeast, our analysis implies that the annual burn area that results from these wildfires is modulated by environmental conditions and therefore future burn-area totals are unlikely to be independent of climate change. Thus, a singular focus on human-ignitions as the driver of southeast wildfire occurrence (expanding season, number of wildfires, fuel moisture where wildfires occur, *etc.*) and ignoring climate change, is not be the most productive way to prepare for the future.

2.3.6   PM2.5 emission estimates by region and ignition type

Wildfires emit particulate matter pollution, which can degrade local and regional air quality (Baker et al. 2016; Brey et al. 2018b; Brey and Fischer 2016; Ford et al. 2017; Jaffe et al. 2008; Lassman et al. 2017; Saide et al. 2015; Val Martin et al. 2015; Wiedinmyer et al. 2006). In the present study, the potential air quality impacts of human- and lightning-ignited wildfires are estimated based on the $PM_{2.5}$ emitted by

these wildfires. The PM$_{2.5}$ emissions presented here were estimated for each FPA FOD wildfire between 2002 and 2015 using the FINN model framework, described in detail in section 2.2.6.
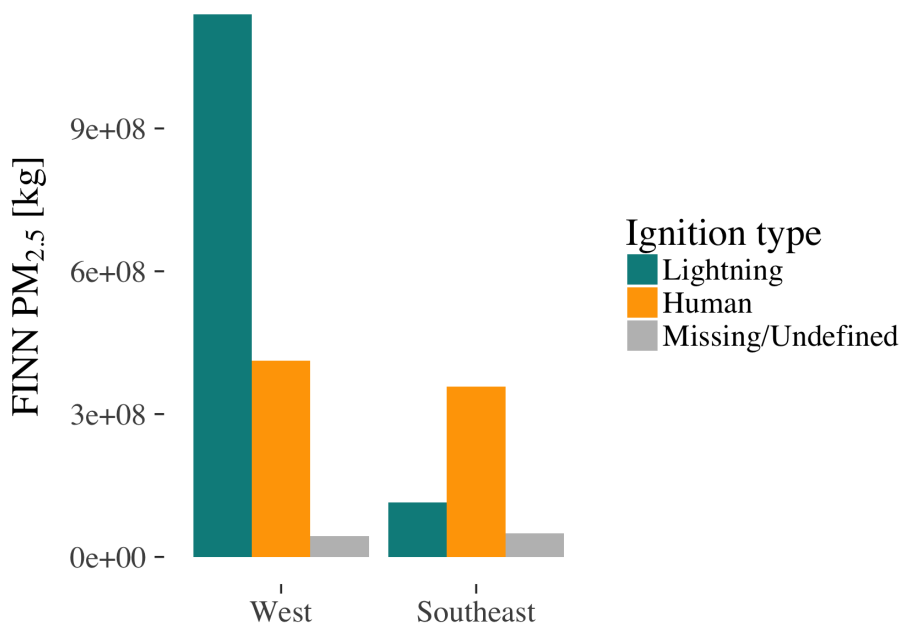

F$_{IG}$. 2.9. The total PM$_{2.5}$ emissions from FPA FOD wildfires between 2002 and 2015 in the west and southeast color coded by ignition type. PM$_{2.5}$ emissions were made using the FINN model framework. FPA FOD wildfires between 2002 and 2015 and all months were used (n=1,086,712), these are the years the two data sources overlap. Only FINN emissions derived from FPA FOD wildfires are shown.

The bars in Figure 2.9 show the total estimated PM$_{2.5}$ emissions of FPA FOD wildfires within the west and southeast regions between 2002 and 2015. The blue and orange bars in Figure 2.9 show the cumulative PM$_{2.5}$ emissions that result from lightning- and human-ignited wildfires respectively. The ignition-type that accounts for the most PM$_{2.5}$ is the same as the ignition type responsible for the majority of burn area within a given region (Figure 2.4). There is more wildfire PM$_{2.5}$ emitted in the west than the southeast between 2002 and 2015. However, there are similar amounts of PM$_{2.5}$ that are a result of human-ignited wildfires in both regions. Although southeast emissions are lower than the west, it is clear that southeast wildfires, although on average smaller, in aggregate emit substantial amounts of PM$_{2.5}$ which have the potential to significantly impact air quality. The PM$_{2.5}$ emissions presented in this section are an incomplete account of PM$_{2.5}$ in both regions since prescribed fire, agricultural burning, and fires on private land are omitted from the FPA FOD.

2.3.7   Ignition types for wildfires leading to U.S. smoke air quality forecasts

This section examines the air quality impacts of FPA FOD wildfires through a different lens, by investigating how often FPA FOD wildfires can be linked to NWS smoke forecasts. This provides a second perspective on the possible air quality impacts of wildfires of different ignition types. We identified the wildfires that were co-located in space and time with HYSPLIT points (human analyzed wildfires used to initiate NWS smoke forecast described in section 2.2.7). We consider HYSPLIT points and FPA FOD wildfires to be co-located, or paired, when two conditions are met. 1) The HYSPLIT point must have been analyzed on the day before or up to seven days after the wildfire discovery date documented in the FPA FOD. 2) The distance between the HYSPLIT point and the FPA FOD wildfire location must be no greater than 10 km. These combined criteria are very strict and most likely underestimate the number of HYSPLIT point detections paired to large FPA FOD wildfires and overestimate the number of HYSPLIT point detections paired to small FPA FOD wildfires. FPA FOD data between 2007 and 2015 was used for pairing, as these are the years FPA FOD and HYSPLIT point data overlap. When we adjust our pairing criteria by changing the distance and time requirements for FPA FOD wildfires and HYSPLIT points to be considered co-located, the total number of wildfires paired changes, but the patterns shown in Figure 10 and presented in the following paragraphs do not change significantly.

The majority of FPA FOD wildfires are not paired with HYSPLIT points. Between 2007 and 2015, only 5.5% of FPA FOD wildfires (5.2% in the west, 5.8% in the southeast) are co-located with HYSPLIT points. The low number of FPA FOD wildfires paired to HYSPLIT points is not surprising given the abundance of very small (<10 acres) wildfires documented in the FPA FOD and the challenge of detecting small wildfires via satellite. The similar percent of wildfires paired in the southeast and west is surprising, given that wildfires in the west are on average larger than wildfires in the southeast. It is likely that our methods underestimate the number of co-located HYSPLIT points for very large FPA FOD wildfires since large wildfires can last longer than the allowable time period for pairing (1 day before and 7 days after the wildfire discovery date). Additionally, the wildfire locations provided in the FPA FOD data are not necessary burn scar centroids, and wildfire burn scars are not always circular, which compromises our method's ability to pair all appropriate HYSPLIT points to large wildfires.
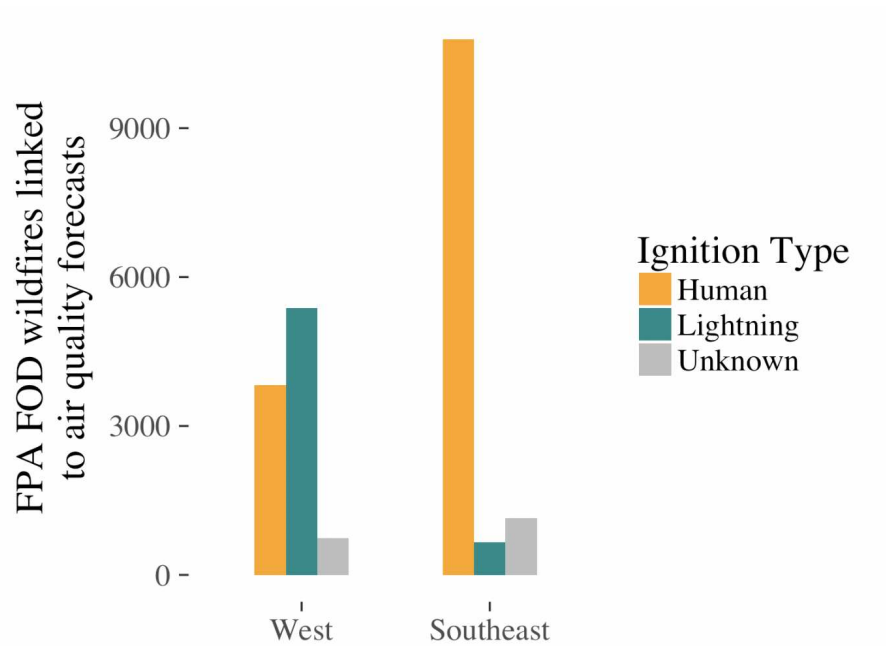
FIG. 2.10. Count of FPA FOD wildfires associated with NESDIS HMS analyzed HYSPLIT points used to initiate the NWS smoke forecast, occurring in the west and southeast regions between 2007 and 2015 (the years these data overlap). The colors indicate the cause of the FPA FOD wildfire. FPA FOD wildfires that were not co-located with HYSPLIT points have been excluded from this plot. There are 9,938 wildfires in the West and 12,588 in the Southeast associated with HYSPLIT points.

2.4 CONCLUSIONS

We use the Fire Program Analysis Wildfire Occurrence Data (FPA FOD) to contrast the environmental conditions, meteorological drivers and air quality impacts of human- versus lightning-ignited wildfires in the southeast and western U.S. We find that the proportion of wildfires that are started by humans is higher in the southeast U.S. than in the west, though the seasonality of when these wildfires occur is also different (*i.e.*, the wildfires in the southeast are bimodal with more occurrences in the spring and fall months, whereas the majority of wildfires in the west occur during July/August). We show that there are larger contrasts in 1000-hour fuel-moisture between ecoregions than between ignition types, which implies that both ignition types are similarly constrained by fuel-moisture within a given ecoregion. Presently, both human- and lightning-ignited wildfire burn area are anti-correlated with total annual precipitation and will likely react similarly to future changes in precipitation (assuming a stationary relationship between precipitation and burn area). Between 1992 and 2015 humans were the dominant source of wildfire ignitions in the southeast U.S. However, the annual burn area

of these wildfires is still linked to environmental conditions that allow fuels to ignite and wildfires to spread. Thus, climate change, not just human-ignited wildfires, will be an important driver of future wildfire activity and the resulting air quality impacts in the southeast U.S. The same is true for the west, where summertime burn area for both ignition types is greater in warmer, drier years. On average, wildfires in the southeast are smaller than in the west. However, these small wildfires significantly impact southeast air quality because 1) there is a large number of southeast wildfires associated with National Weather Service air quality smoke forecasts, and 2) total $PM_{2.5}$ emissions from human-ignited wildfires in the southeast are similar to the total $PM_{2.5}$ emissions from human-ignited wildfires in the west.

CHAPTER 3

EXAMINING THE FUTURE SPREAD IN ENVIRONMENTAL CONDITIONS THAT DRIVE WESTERN UNITED STATES SUMMERTIME WILDFIRE BURN AREA[2]

3.1 INTRODUCTION

Western United States (U.S.) wildfires can impact atmospheric composition on local, synoptic, and continental scales (Baker et al. 2016; Brey et al. 2018a; Brey and Fischer 2016; Ford et al. 2017; Lassman et al. 2017; Saide et al. 2015; Val Martin et al. 2015; Wiedinmyer et al. 2006). The air quality impacts of Wildfire smoke can be especially pronounced in the western U.S. (McClure and Jaffe 2018c; O'Dell et al. 2019), where the majority of particulate matter with diameters less than 2.5 microns ($PM_{2.5}$) National Ambient Air Quality Standards (NAAQS) exceedances can be attributed to wildfire smoke (Liu et al. 2016a). A growing body of research has demonstrated that ozone, a gas phase criteria pollutant in the U.S., can also be enhanced in the presence of wildfire smoke (Brey and Fischer 2016; Jaffe et al. 2013; Lu et al. 2016; McClure and Jaffe 2018b). The deteriorated air quality due to wildfire smoke can impact public health (Rappold et al. 2011; Delfino et al. 2009; Fisk and Chan 2017; Lipner et al. 2019; Gan et al. 2017). Recent work has also begun to identify the impacts of smoke from wildfires on radiation (*e.g.* McKendry et al. (2019)) and terrestrial productivity (*e.g.* Yue and Unger (2018)).

The cost to suppress wildfires has increased in recent years, exceeding 50% of the United States Forest Service (USFS) annual budget for the first time in 2015. In 1995, fire accounted for only 16% of the budget (USFS 2015). These increased costs come at the expense of other USFS priorities, including recreation and restoration. Recent western U.S. wildfires have had large impacts on people, property, and local economies. The 2018 Camp Fire in Butte county California resulted in 86 civilian deaths, the destruction of much of Paradise California, and 16.5 billion USD in total losses, making it the most expensive natural disaster in 2018 (Reyes-Velarde 2019; McBride 2018; Almukhtar et al. 2018). Stand-replacing wildfires can alter carbon balance by turning wildlands into a source for atmospheric carbon (Kashian et al. 2006). In the long term, if stands grow back, the net impact on atmospheric carbon would be close to zero (Kashian et al. 2006), however, as the climate warms, shifts in vegetation types, such as trees to shrubs/grasslands, become more likely (Parks et al. 2018). Given this suite of impacts,

---

[2]Adapted from a manuscript submitted to *AGU Earth's Future,* June 2019

there is a need to understand what drives year-to-year wildfire activity, how those drivers may change in the future, and how future burn area and total emissions may respond.

Many factors contribute to variations in western U.S. wildfire activity including natural climate variability, climate change, and land management (Abatzoglou and Williams 2016; Pechony and Shindell 2010; Short 2015; Westerling et al. 2006, 2014). Variability in long-term climate and year-to-year weather can influence the availability and flammability of fuel (Littell et al. 2009, 2016). Studies have shown that western U.S. wildfire burn area has increased since the 1980s as the western U.S. has warmed (*e.g.* Westerling et al. (2006)).

From an energy perspective, temperature alone is not an ideal measure of wildfire potential. At the scale of individual wildfires, fire spread is a series of ignitions where heat from the fire raises fuels to ignition temperature (Rothermel 1983; Simms and Law 1967). Before ignition of additional fuel can occur, water must be evaporated from the fuel, which comes at the cost of the latent heat of vaporization (2257 J/g); thus the rate of fire spread depends on fuel-moisture content (Bradshaw et al. 1984; Fosberg et al. 1971; Rothermel 1972; Simms and Law 1967; Stocks et al. 1989). Wildfires consume both live and dead fuels (Jolly and Johnson 2018) and are heavily influenced by fuel moisture content (Rothermel 1972; Simms and Law 1967). Living plants can reduce fire spread if they are not water stressed; as they typically have moisture content values an order of magnitude higher than dead fuels (Cohen et al. 1990). Thus, environmental variables that modulate moisture available to plants may be more physically relevant to live fuels flammability than atmospheric temperature. In contrast, dead fuels, such as litter, passively respond to ambient atmospheric conditions.

Given the myriad of environmental variables that have been identified as possible drivers of western U.S. wildfires. Statistical models designed to explain observed variance or predict future fire activity often leverage correlations between wildfire burn area, temperature, and other environmental variables influenced by temperature. On seasonal-to-annual timescales, western U.S. wildfire activity (*e.g.* burn area) can be highly correlated with temperature, vapor pressure deficit (VPD), hot-dry-windy days, soil moisture, relative humidity, and precipitation (Abatzoglou and Williams 2016; Brey et al. 2018a; Forkel et al. 2017; Kloster and Lasslop 2017; Littell et al. 2009; Parks et al. 2018; Park Williams et al. 2012, 2017; Pechony and Shindell 2010; Rothermel 1983; Short 2015; Srock et al. 2018; Westerling et al. 2006, 2014; Yue et al. 2013b). Different environmental variables have been shown to be important in different regions of the western U.S. For example, wildfire burn area in the southwest U.S. has been shown to be related to drought at varying antecedent timescales (Park Williams et al. 2015). Littell et al.

(2009) showed that 1916-2003 wildfire burn area in the southwest was driven by fuel availability, the previous year's precipitation, and least of all in their analysis, the conditions of the present year. Abatzoglou and Williams (2016) show that metrics of aridity (*e.g.* Vapor Pressure Deficit) are well correlated with western U.S. forested burn area. Park Williams et al. (2015) also show that VPD is well correlated with western U.S. wildfire burn area. Holden et al. (2018) show that variability in wildfire season precipitation (*e.g.* days with rain > 2.54 mm) control summer burn area and strongly influence VPD. Westerling (2016) show that warming and the timing of spring snowmelt can modulate burn area. Jensen et al. (2018) found that in some areas wildfires become more likely following months with high soil moisture values. Many studies that leverage these observed relationships predict significant increases in western U.S. wildfire activity under climate change (*e.g.* Liu et al. (2016b); Flannigan et al. (2009); McKENZIE et al. (2004); Westerling et al. (2011); Yue et al. (2013b); Spracklen et al. (2009)). Temperature is physically related to each environmental predictor listed here, however, despite being physically related, relationships between temperature and other environmental predictors can be non-linear, resulting in differing and less certain future changes than temperature, as simulated by global climate models.

One of the acknowledged weaknesses associated with statistical model based approaches to predicting future wildfire activity is the assumption that historical relationships will remain stationary in the future. Using the Northern Rockies as an example, Higuera et al. (2015) showed that these relationships change over time. In addition, due to internal variability of the climate system and anthropogenic-caused climate change, the climate of the 21st century will include mean and extreme states of environmental variables that were not observed in the historical period (Collins et al. 2013; Mora et al. 2013). Further, Parks et al. (2018) showed that there may be a tipping point where vegetation shifts from, *e.g.*, forests to shrubland/grasslands as the climate changes, providing evidence that fuels will not remain the same.

The overarching goal of this work is to understand what environmental factors contribute to the uncertainty of 21st century western U.S. wildfire activity quantified by burn area, a metric proportional but not perfectly correlated to air quality impacts (Brey et al. (2018a)). While previous studies have shown there are many factors that contribute to this uncertainty (*e.g.* emission factors, changing wildland-urban interface, sources and abundances of ignition sources, changing fuel loads), our work quantifies the individual environmental conditions contributions to future burn area estimates, and

compare this spread to scenario and intermodel spread across the mountainous western U.S. ecoregions. These ecoregions contain much of the forested mountains of the western U.S., highest carbon emission factors, and fuel loading (kg fuel m^{-2}) (Abatzoglou and Rupp 2017; Urbanski et al. 2018; van der Werf et al. 2010). These ecoregions were the largest emitters of smoke PM$_{2.5}$ in the western U.S. between 2003 and 2015 (Urbanski et al. 2018). Thus, we focus on these regions for this analysis. Our methods do not account for changes in fuel loading, so we continue our analysis motivated by the assumption that these regions will continue to be characterized by large fuel loading and emission factors in future decades.

One of the novel contributions of this work is that we employ an objective method for choosing environmental variables that explain historical wildfire variability, allowing us to build on the relationships observed by previous studies while putting forth a new method for identifying key variables. We define environmental conditions previously identified to be important for wildfire occurrence as "candidate variables". The candidate variables used in this work include seasonal temperature, wind speed, precipitation, evaporation, relative humidity, and root zone soil moisture. These variables 1) have been demonstrated to impact wildfire occurrence by other studies and/or 2) are related to or are an indicator of the moisture budgets that are relevant for fuel availability, flammability, or plant stress, and 3) are available in both historical reanalysis datasets and CMIP5 future-projection simulation output, which enables the training of statistical models from historical information and use of these models for forward projections. For a given ecoregion, we objectively select the variables that best explain observed variability using a custom implementation of Lasso regression. This approach allows for us to learn from the data rather than confirm a hypothesis of a specific variables importance. Because temperature is often described as important, and is physically related to all of the candidate variables, our work examines how estimates of future wildfire change when temperature is and is not used as a predictor using the following approach.

(1) We use Lasso regression to objectively identify the environmental conditions that best explain observed variance in interannual summertime (June, July, August) large wildfire (> 1000 acres) burn area for western U.S. ecoregions. We focus on these months as they have the greatest burn area in recent decades (Brey et al. 2018a). This is done individually for ecoregions, to allow for varying relationships to be accounted for.

(2) We use CMIP5 model output to quantify the mean change and future spread in the environmental variables that best explain historical variance for each fire-prone ecoregion. We contrast the trends and spread between models and discuss what the changes may imply for future burn area.

(3) Finally, we combine historical relationships between burn area and select candidate variables with CMIP5 output to examine the spread in future burn area as forecast by the objectively selected variables and compare this to the spread between RCP 4.5 and 8.5 scenarios. We contrast how future estimates driven by our statistical model (trend, variability, *etc*.) compare to future estimates driven by changes in temperature alone.

This work shows that the predicted change in burn area is extremely sensitive to what environmental predictors are chosen to drive burn area.

3.2 DATA AND METHODS

3.2.1 *ECMWF ERA-Interim Reanalysis*

For historical data on environmental variables, we use European Center for Medium-Range Weather Forecasts (ECMWF) ERA-Interim monthly reanalysis fields (followed by ERA-Interim specific abbreviations) (Dee et al. 2011) of temperature (t2m), total precipitation (tp), wind speed (si10), relative humidity (calculated using dew point temperature, surface pressure, and Tetens formula), soil moisture (mrso), and evaporation (e). All data were downloaded at the native horizontal grid spacing of 0.75° x 0.75°. To make the ERA-Interim output directly comparable to CMIP5 output (described next), variables were bilinearly re-gridded to a common grid with 2°x2.5° degree grid spacing Schulzweida (2019). Where necessary accumulation units (*e.g.* mm day^{-1}) were converted to flux units (*e.g.* $\text{kg m}^{-2}\text{ s}^{-1}$) to match the units used by CMIP5 models. A total root zone (3 meter depth) soil moisture field was created by converting volumetric soil water in layers 1 through 4 (total depth of 2.89 m) to kg m^{-2} by assuming the density of all water content to be 1000 kg m^{-3}.

3.2.2 *CMIP5*

To estimate the future change and spread in wildfire-relevant environmental predictors, we use output from the Phase 5 of the Coupled Model Intercomparison Project (CMIP5) (Taylor et al. 2012). We use the following monthly variables (followed by CMIP5 specific abbreviations): near surface temperature (tas), precipitation (pr), Near-Surface Wind Speed (sfcWind), near-surface relative humidity (hurs), soil moisture content (mrlsl), and evaporation (evspsbl). Intermodel spread between models is

quantified using the r1i1p1 ensemble member for each climate model. We are interested in quantifying intermodel spread and scenario spread, which can be accomplished with a single ensemble member. Though not as comprehensive as using multiple ensemble members, the use of a single ensemble member still allows our methods to account for internal variability of the climate system. Scenario spread is quantified by examining output from Representative Concentration Pathways (RCP) 4.5 and 8.5, scenarios that correspond to different greenhouse gas climate forcing (W m^{-2}) scenarios for the 21st century. RCP 4.5 represents an emission scenario where greenhouse gas emissions begin to slow in the mid 21st century, RCP 8.5 represents a scenario where emissions continue to grow throughout. In order to bias-correct CMIP5 data using the reanalysis data, historical (pre-2006) CMIP5 data were also downloaded. The data from all models were regridded to the GFDL-CM3 2°x2.5o° degree grid using bilinear-interpolation for atmospheric domain variables and distance-weighted average remapping for land domain variables (Schulzweida 2019). Data for this work was downloaded interactively from the Earth System Grid Federation (https://esgf-node.llnl.gov/search/cmip5/).

While accurate soil properties are required to accurately simulate soil-water content and implications for seasonal drought stress (Peterman et al. 2014) the soil schemes in the CMIP5 models vary substantially (Berg et al. 2017). For example, the active soil depth varies between 3 and 42 meters. Soil moisture by layer (mrlsl) in the CMIP5 archives include water in all phases. The different active soil moisture depths make intermodel comparison of total soil moisture impractical, so instead, we regrid the soil moisture to include only the top 2.89 meters, which is the part of the soil column most accessible to ecosystems (root zone soil moisture, regridded down to 2.89 m to match ERA-Interim soil moisture depth); this has been used to compare models in other studies (*e.g.* Berg et al. (2017)). A depth of 2.89 is close to the maximum soil depth of several models, which is at a minimum 3 meters. The hydraulic properties of soils will also vary among models, which will partially contribute to differences seen between them. The vertical interpolation is required due to the significantly varying depths and layers output in CMIP5 coupled models. We convert zeros to nans. Due to the different ways CMIP5 models represent soil moisture processes, integrated water in the top 3 m does not have the same meaning in terms of drought stress across models.

3.2.3 *Monitoring Trends in Burn Severity (MTBS)*

This work quantifies interannual variability of wildfire burn area using the wildfires documented in the Monitoring Trends in Burn Severity database (MTBS). MTBS is a multiagency program with the

goal of consistently mapping the burn area and severity of large wildfires in the U.S. in order to monitor the effectiveness of the National Fire Plan and Healthy Forests Restoration Act (Eidenshink et al. 2007). These data include wildfire burn area for the western U.S. between 1984 and 2016 for wildfires larger than 1000 acres (404 ha) (Eidenshink et al. 2007). Wildfires larger than 1000 acres account for the majority of burn area in the western U.S. (Brey et al. 2018a). MTBS uses the Normalized Burn Ratio (dNBR) from Landsat to detect fire burn scars and severity (Eidenshink et al. 2007). MTBS data are largely divided into mapping zones based on Bailey's Ecological Sections (Bailey 1983), but this mapping is slightly altered to account for meaningful administrative boundaries (Eidenshink et al. 2007). Though MTBS assesses the severity of burn area, our work makes no distinction between severities to create an aggregate burn area. MTBS will generally report less burn area than totals tabulated by incident reports (Eidenshink et al. 2007). For example, in 2004 MTBS mapped 7,781,049 acres of wildfires while statistics compiled by the National Interagency Coordination Center (NICC) documented 8,097,880 acres (Eidenshink et al. 2007). MTBS maps wildfires and prescribed (Rx) fires. Our work aims to focus on the climate and biophysical controls on fire occurrence, so we exclude Rx fires from our analysis, and from here forward we refer to the remaining types in MTBS as 'wildfires' (similar to Finco et al. (2012)). One of the fire types in MTBS is "unknown". There is a possibility that some of these fires were prescribed; however, MTBS documents only one documented Rx fire in the western U.S. that occurs in the months June, July, or August, the months of interest in this analysis, so we expect prescribed fires to make up a small amount of the burn area contributed by the unknown fire type category. The 33 years of wildfire data in MTBS provides a longer time period than other incident based datasets that include additional small wildfires (*e.g.* FPA FOD developed by Short (2014)), which grants analysis of trends and variability in wildfire occurrence additional validity. MTBS data were downloaded as Fire Occurrence Data (point locations rather than burn scar shapes) from https://www.mtbs.gov/direct-download. These fire data where then assigned to Baileys divisions to create monthly burn area totals aggregated by areas with similar geographic and ecological characteristics.

3.2.4   *Bailey's Ecoregions*

To account for how environmental drivers of wildfires may differ between regions, we aggregate MTBS burn area data, ERA-Interim reanalysis, and CMIP5 model output by Bailey's ecoregions, which map the regional extent of ecosystems in the U.S. Bailey's ecoregions include four levels, with each subsequent level containing smaller, more narrowly defined ecosystems. We choose to use the "divisions" level, a subset of the "domain" (most coarse) level. These nearly contiguous divisions allow

us to aggregate wildfires that occur in areas of similar temperature and precipitation. The Temperate Steppe, Marine, and Mediterranean Regime Mountain divisions (shown in 3.1) contain much of the forested mountains of the western U.S. (Abatzoglou and Williams 2016), and correspondingly, some of the largest and most destructive wildfires, as well as the highest $PM_{2.5}$ emissions between 2003 and 2015 (Urbanski et al. 2018). Thus, our analysis focuses on these three ecoregions.

MTBS data were downloaded as shapefiles from https://www.fs.fed.us/rm/ecoregions/products/map-ecoregions-united-states/. The divisions (hereafter referred to as "ecoregions" used in this work are shown in Figure 3.1.
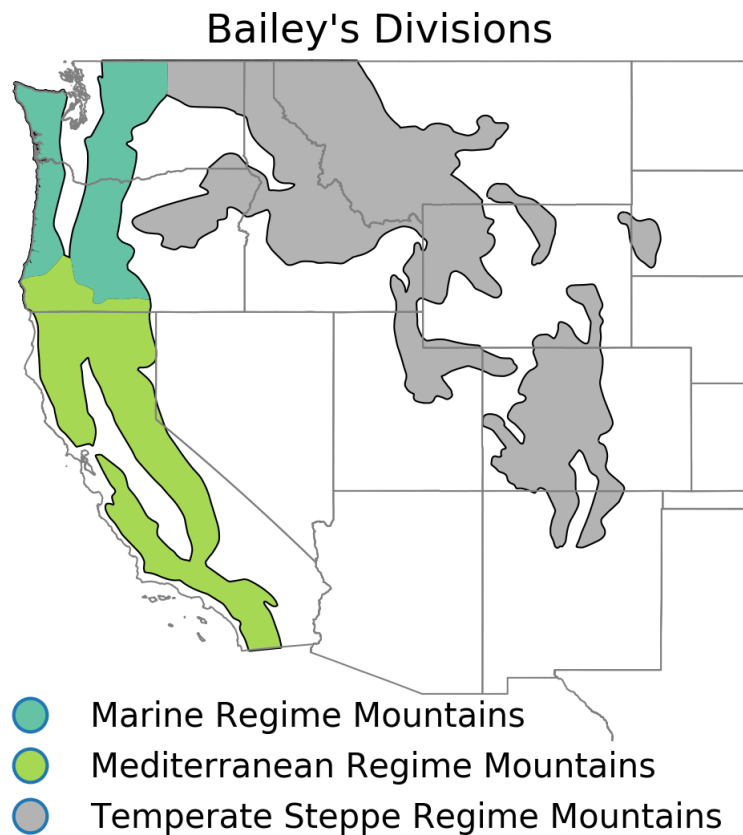

FIG. 3.1. The western U.S. Bailey's divisions analysed in this work (hereafter referred to as ecoregions). The Marine regime mountains in this work excludes portions of that region above 50 degrees latitude.

3.2.5 *Wildfire predictor identification and regularization with Lasso Regression*

We use the Least Absolute Shrinkage and Selection Operator (hereafter Lasso) regression to create regularized linear models for each of the three ecoregions in the western US shown in Figure 3.1. Lasso is a regression method that performs regularization and variable selection (Tibshirani 1996). The

Lasso was developed to increase the prediction accuracy of ordinary least squares (OLS) regression by reducing the variance of predicted values (Tibshirani 1996). The Lasso accomplishes high prediction accuracy and lower variance through a constraint on the sum of the absolute values of the model parameters, which results in some predictor coefficients being set to exactly zero. When the regularization term α is zero, the Lasso is the same as OLS. The higher the value of α, the more regularized (simpler the model is forced to be to avoid overfitting) the regression model becomes, and more predictors coefficients are set to exactly zero. Lasso has been demonstrated to make a good compromise between model complexity and model performance (Loukina et al. 2015). The Lasso can select multiple correlated predictors. For example, if dry spring seasons are correlated to dry summer seasons, and both are good predictors of wildfires, using both gives the desired insights and variance explained needed for this work.

$$\frac{1}{N}\sum_i (y_i - \mathbf{w}^T\mathbf{X})^2 + \alpha||\mathbf{w}||_1 \qquad (3.1)$$

Equation 3.1 shows the Lasso cost function, the value of which is minimized in the Lasso regression. y is the target function, which in this work is summer burn area. $\mathbf{X}$ is a matrix of features where each column is an environmental predictor. N is the number of examples (rows in $\mathbf{X}$), in our work this corresponds to 33 years of wildfire data and environmental predictors (1984-2016). $\mathbf{w}$ is a vector of coefficients that corresponds to the columns in $\mathbf{X}$. α determines the degree of regularization for the model. We find the optimal value of α using leave one-out-cross validation. The L1 norm ($||\cdot||_1$) is not differentiable, so the Lasso does not have an analytical solution and must be solved numerically. We implement the Lasso using the LassoCV method from scikit-learn (Pedregosa et al. 2011). The optimization for the Lasso is solved with coordinate descent.

We implement the so-called "relaxed Lasso", where Lasso regression is used twice (Meinshausen 2007). This is done in two steps. First, we use Lasso for variable selection. This is done by leaving out a year of data (*e.g.* 1984), fit a Lasso regression using leave-one-out cross validation on the remaining years (1985-2016) and recording what variables receive a non-zero coefficient. We repeat this process for each year. We discard the 50% of features that were selected (non-zero coefficient) the least often. This step helps us identify what proportion of the time a given features is selected, which gives a measure of how robust that feature is to random chance in the optimization and noise in the yearly data. Second, we fit the Lasso regression using leave-one-out cross validation using the 50% of the original

features that were selected most frequently in step 1. This is where the final optimal value of α is selected, which is the value that provides the lowest mean square error on years left out. Because feature selection has already been performed once, the values of α in the second implementation are smaller than in the first (*i.e.* less regularization). This two step process allows us to quantify how robust the selected features are to noise in the data.

In our use of Lasso, environmental variables from ERA-Interim reanalysis are used as predictor features. These features include all of the candidate environmental variables listed in the introduction, where the average value for each feature is taken for the three different seasons. ERA-Interim data are spatially subset to include only the grid boxes that overlap a given ecoregion's geographic extent. This results in larger ecoregions being represented by more grid boxes than smaller ecoregions. We take the temporal average of overlapping features for three seasons, winter (months 11, 12, 1, 2), spring (months 3,4,5) and summer (months 6,7,8). This temporal averaging is a necessary step to allow for relationships between antecedent environmental variables and summer burn area observed in other studies to be accounted for in this analysis. Between the antecedent seasons and summer temporal means, and 6 variables, there are a total of 18 features (columns of **X**) available to explain 33 years of variance in summer wildfire data (y).

The underlying assumption of this work is that year-to-year variability relationships can be leveraged to estimate long-term changes in burn area. Due to the limited quantity of wildfire data, we set up our regression to explain variance in year-to-year wildfire burn area, rather than the trend in features or wildfire burn area. To avoid a spurious model where the target function and predictors are well correlated because of a shared trend, we detrend the environmental variables and the burn area time series before fitting the regression. This ensures that the regression predicts year-to-year variability and that correlations between burned area and features are due to interannual co-variability. Regressions are performed independently for each ecoregion. Similar to other work (*e.g.* Littell et al. (2009); Park Williams et al. (2015)), the summer burned area was $\log_{10}$ transformed to account for the exponential distribution of annual burned area and to mitigate heteroscedasticity. Heteroscedasticity results from year-to-year quantities that span orders of magnitude, increasing over time, which results in variance that is not stationary in time. Without taking the $\log_{10}$ transformation of summer burn area, some ecoregions exhibited upward trends in model residuals. In addition, we observe better linear model performance (R^2 values) when summer burn area is $\log_{10}$ transformed. Because the $\log_{10}$ of zero is undefined, we set years with zero acres of burn area to 1. This only occurs in the Marine Regime

Mountains. In order to make comparing environmental variances with different units easier, we standardize each feature (μ=0, σ=1) for the historical training period (1984-2016) such that the units of linear model regression coefficients are the same ($\log_{10}$(burn area) σ^{-1}).

3.2.6   *Quantification of Spread In Key Wildfire Predictors*

We use CMIP5 model output to show how environmental variables (*i.e.* features) selected by the Lasso could change in the future. We examine changes in year-to-year variability as well as changes due to long term trends. Similar to Park Williams et al. (2012, 2015), we bias-correct CMIP5 modeled historical and future projections interannual standard deviation and mean value between 1984 and 2016 to match those estimated by the ERA-Interim reanalysis fields. This was done for each feature predicted by each CMIP5 model individually using the following procedure. 1) The linear trend is removed from the 1984-2016 CMIP5 feature times series (*e.g.* summer precipitation for model ACCESS1-0). 2) The standard deviation, which after detrending is due entirely to year-to-year variability, is calculated for the detrended 1984-2016 time series. 2) The detrended 1984-2100 CMIP5 time series is multiplied by the ratio of the detrended ERA-Interim feature standard deviation over the detrended CMIP5 feature 1984-2016 standard deviation. At the completion of this step, the CMIP5 feature standard deviation for 1984-2016 matches ERA-Interim. 3) The 1984-2016 linear trend is added back to the CMIP5 feature. 4) To correct the CMIP5 model bias offset, the mean value of the time series from 1984-2016 is replaced by the ERA-Interim mean for the same years.

Differences in modeled and observed standard deviations and trends suggest that models are not simulating what is estimated by ERA-Interim. Scaling these data can also inflate confidence in models as they have been forced to look more similar than the original output. However, we choose to scale the variance in this work because it eases the interpretation of the changes in environmental predictor values in the following ways. 1) It allows for a straightforward interpretation of changes observed by models from the historical period to the simulated future. 2) It gives us the ability to observe how the spread in a variable changes due to changes in the range of year-to-year values. 3) It ensures the spread in the estimates of future wildfire activity estimated by CMIP5 output using the regression models is not artificially narrow due to a small variance in CMIP5 model variables. Finally, bias-corrected CMIP5 features are standardized in the same way as the ERA-Interim features used to train the Lasso, by subtracting the 1984-2016 ERA-Interim mean and dividing by the standard deviation. This adds additional interpretability to the results. For example, if a CMIP5 model feature value for summer precipitation

in the year 2019 was zero that would mean that CMIP5 model simulated summer 2019 precipitation to be the same as the mean historical value.

Because Lasso regression identifies different predictors for different ecoregions, the number of CMIP5 models with all required variables available varies by ecoregion and type of regression. For example, fewer CMIP5 models save out near-surface relative humidity than near-surface temperature. As a result, an ecoregion with only temperature features selected by the relaxed Lasso would have more CMIP5 models with output available. This makes comparing spread in variables between ecoregions challenging, as the spread are partially attributed to the different number of CMIP5 models. However, within an ecoregion, there are always a consistent number of models representing each feature, as we require a given CMIP5 model to have output for every feature selected by the Lasso in order to be used to examine spread or project future wildfire burn area.

3.3 RESULTS

3.3.1 *Historical relationships between wildfire burn area and predictors*

As discussed in the introduction, there are many environmental variables that have been shown to be correlated with wildfires at different antecedent and concurrent timescales. Our work objectively identifies the environmental variables that offer the best prediction of total summer (June, July and August) wildfire burn area for mountainous western U.S. ecoregions and compares the variance explained by the regression fit with access to all candidate variables (hereafter "Lasso-All") to the regression fit with only temperature (hereafter "Lasso-Temperature"). The candidate variables are listed in the methods, they include, temperature, wind speed, precipitation, evaporation, relative humidity, and root zone soil moisture, averaged for Winter (11,12,1,2), Spring (3,4,5), and Summer (6,7,8) months. Figure 3.2 presents the historical variance explained in summertime $\log_{10}$ wildfire burn area within western U.S. ecoregions for each regression. The Lasso-Temperature regression explains more than half of the observed variability in the Temperate Steppe Regime Mountains; in this region temperature alone is able to explain a similar amount of variance as the Lasso-All regression. The variance explained by ordinary least squares would be higher than the values presented in Figure 3.2.
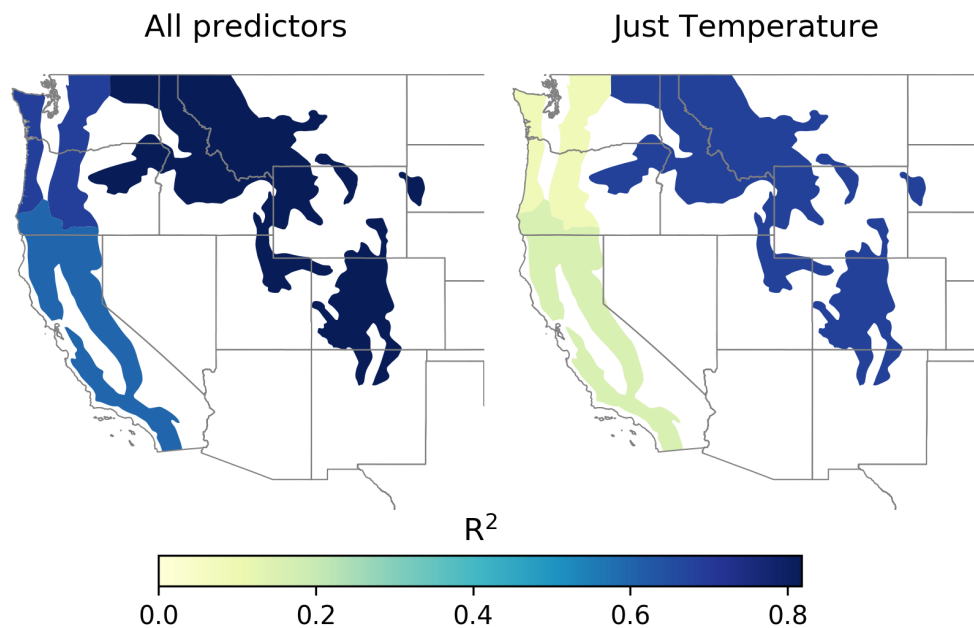
Fig. 3.2. Historical variance in year-to-year June-August log10 wildfire burn area explained by Lasso regressions for the presented western U.S. ecoregions using ERA-Interim features as predictors for the trained regressions. Note, the colorbar scale does not go to 1.0.

There are several results in Figure 3.2 relevant to forward predictions of burn area. The trained regression generally misses the extremes in interannual variability (shown in the Appendix Figures A5 through A5), *i.e.* summers with very high burn area are under-forecast and summers with very low burn area are over-forecast (*i.e.* the burn area forecast lack sharpness). As a result, the year-to-year signals are captured, but the representation of year-to-year extremes are underestimated, an important attribute of these regressions to consider when predicting burn area using CMIP5 model out. Figure 3.2 shows that temperature can explain wildfire activity but that linear combinations of other variables can as well. The variance explained by the left hand panel, the Lasso-All regression, is mostly *not* temperature (shown next), which indicates that temperature is not required to explain historical variability in wildfire burn area.
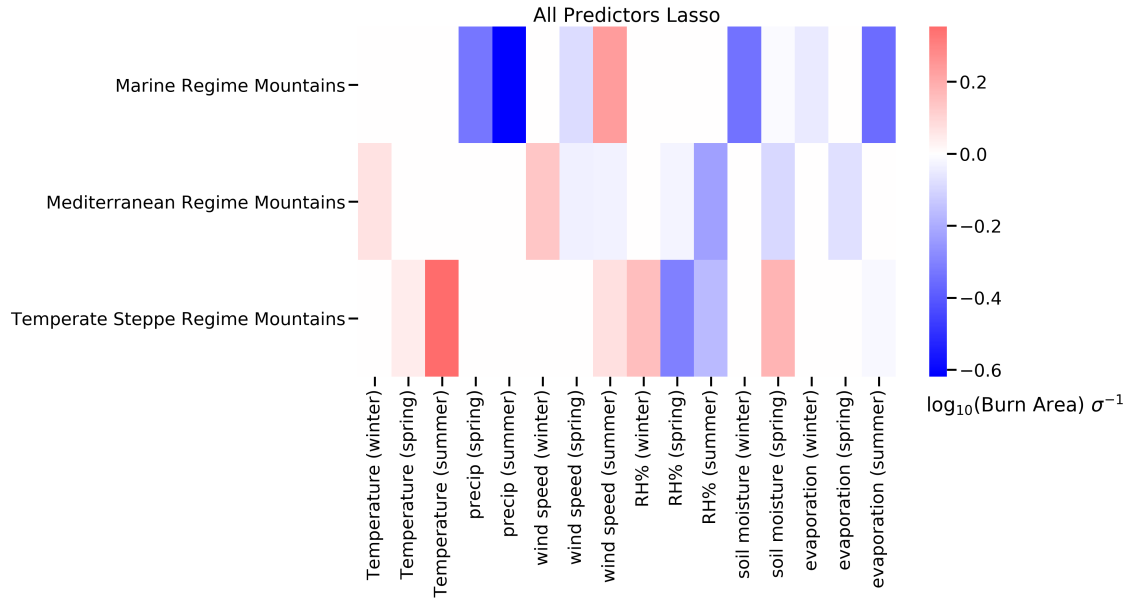
FIG. 3.3. Coefficients from the Lasso-All regression (access to all candidate variables) (horizontal axis) for the three western U.S. ecoregions used in this analysis (vertical axis). Only predictors with a non-zero value in at least one ecoregion are shown along the vertical axis of this plot.

The colors in Figure 3.3 represent the magnitude of coefficients for the Lasso-All regression. The larger the absolute value of the coefficient, the larger impact that variable has on predictions per unit change of that predictor. Coefficients can be compared across ecoregions, and many variables have coefficients equal to zero, indicating that the linear model with the most skill usually only needs a small subset of the available predictors. This plot shows that among the mountainous ecoregions, temperature is selected in the Temperate Steppe (spring and summer) and Mediterranean Regime Mountains (winter). Summer relative humidity is strongly anti-correlated with burn area in every ecoregion where it is selected. Most features have the same sign across all ecoregions where they are used, indicating a consistent relationship between features and burn area across ecoregions. Our results show that surface wind speed is used in all ecoregions though it does not have a large coefficient relative to other variables. Other studies have also noted weak relationships between wind speed and burn area for gridded fire and meteorology data (Brey et al. 2018a; Crevoisier et al. 2007; Yue et al. 2013b), it is possible, that wind speed is acting as a proxy for other wildfire relevant conditions, like the location of seasonal storm tracks. All of these variables are physically related to temperature, but our methods account for this. When temperature is not selected as frequently as other variables, it means that it did not consistently provide as small of an error on unseen data as other variables.

3.3.2   Spread In Wildfire Predictors and Influence on Future Fire Projections

This section focuses on how candidate environmental variables are predicted to change as estimated by CMIP5 models. Figure 4 shows the multimodel mean change in the mean value of each feature from the historical period (1984-2016) to the future period (2070-2099) against the Lasso regression coefficient. The whiskers show the range (min and max) of change estimated by all models. The horizontal axis of panels in Figure 4 shows the value of the Lasso-All candidate variables regression coefficients. The further a candidate variable (colored dot and whiskers) is plotted from the dashed vertical line, the bigger impact that feature has on predictions of future burn area per unit change in that variable (expressed in terms of historical standard deviation). The vertical axis of panels in Figure 4 shows how much a candidate variable is simulated to change from the historical period to the future period. The further a candidate variable is from the dashed horizontal line the more the variable has changed between the historical and future time period as estimated by CMIP5 models. The features in the top right and bottom left quadrants (light red background shading) imply the mean change as estimated by CMIP5 models will result in increased burn area. The majority of features are plotted in the red shaded regions for all ecoregions, indicating future conditions will resemble even more extreme conditions than those of high burn area years from the historical period. The multimodel mean changes in feature values are usually larger for RCP 8.5 than RCP 4.5. Temperature increases in all three seasons for all three ecoregions (rows). In terms of its own historical variability, temperature increases more than any other variable type.
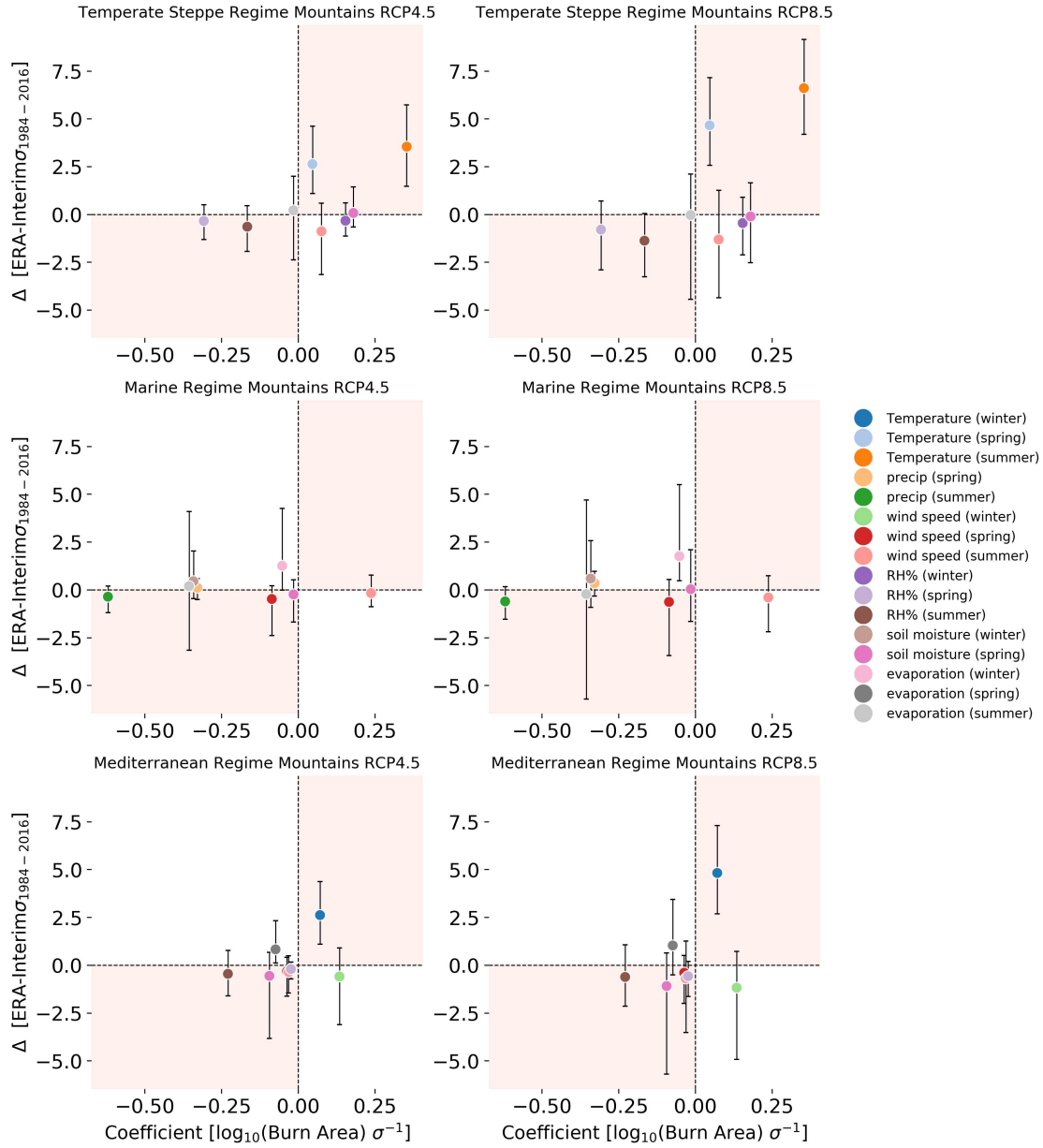
FIG. 3.4. Each panel shows the multimodel mean change in features mean value from 1984-2016 to 2070-2099 (vertical axis) against the Lasso regression coefficient (horizontal axis). The change in feature value is expressed in units of historical σ as estimated by detrended ERA-Interim feature values from 1984-2016. The black whiskers show the total range in changes observed by CMIP5 models. Each row corresponds to an ecoregion and the columns show changes for a given RCP (4.5 on the left, 8.5 on the right). The lower-left and upper-right portions of the plot are shaded light red to indicate that symbols that fall into these quadrants imply changes in variables that correspond to increases in burn area according to the Lasso all regression.

Figure 3.5 shows how the long-term changes simulated for different features vary between models and ecoregions (for variables with non-zero Lasso regression coefficients for the Lasso-All regression). Every CMIP5 model for both RCP scenarios shows an increase in the winter, spring, and summer temperature for the ecoregions with non-zero regression coefficients from the historical period to the future period. The spread in temperature is comparable to other features, but it is unique in that the estimates show only increases, across all models, RCP scenarios, and ecoregions. In contrast, CMIP5 models show both increases and decreases in every other predictor for both RCPs. Summer relative humidity is anti-correlated with summer burn area in each region where it is selected and Figure 3.5 shows the majority of CMIP5 models predict decreased summer relative humidity, but some also show increases. Most models show a decrease in summer precipitation in the Marine Regime Mountains, the variable with the largest coefficient absolute value in that region. Most models estimate decreased relative humidity in all three seasons for the Mediterranean and Temperate Steppe regime mountains. There is disagreement on the sign and magnitude of summer evaporation.
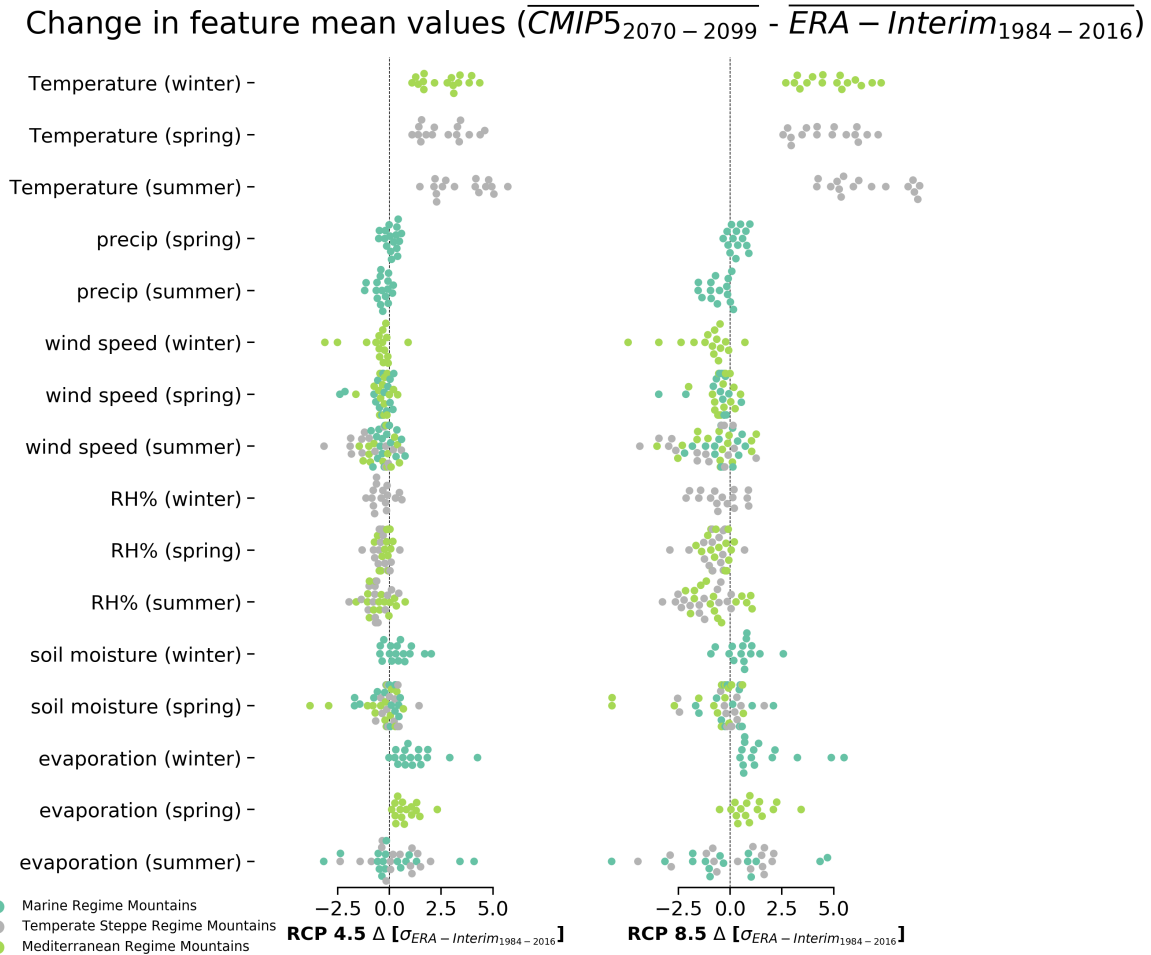
FIG. 3.5. The change in feature mean value (horizontal axis) from 1984-2016 to 2070-2099 shown for different features (rows) and ecoregions (colors). Each dot shows this change estimated by a single CMIP5 model for RCP 4.5 (left) and RCP 8.5 (right). The change in feature value is expressed in units of historical σ as estimated by detrended ERA-Interim feature values from 1984-2016. Values that fall on the vertical line are features where the multi-decade mean value has not changed between the two time periods. For a given ecoregion (color) the numbers of CMIP5 models (dots) is consistent, but can differ between ecoregions. Changes in feature value estimates for a given ecoregion (color) are shown where the absolute value of the Lasso regression coefficient (not shown) for that feature was greater than zero.

3.3.3 Projected wildfire burn area using historical relationships

This section compares the spread in future burn area across RCP scenarios and regressions driven by CMIP5 output. This is done for each CMIP5 model independently, so each CMIP5 model makes an estimate of summer burn area for all years between 1984 and 2099. This is repeated for both RCPs and Lasso regressions. Figure 3.6 shows the CMIP5 ensemble mean burn area estimated by each of these combinations.
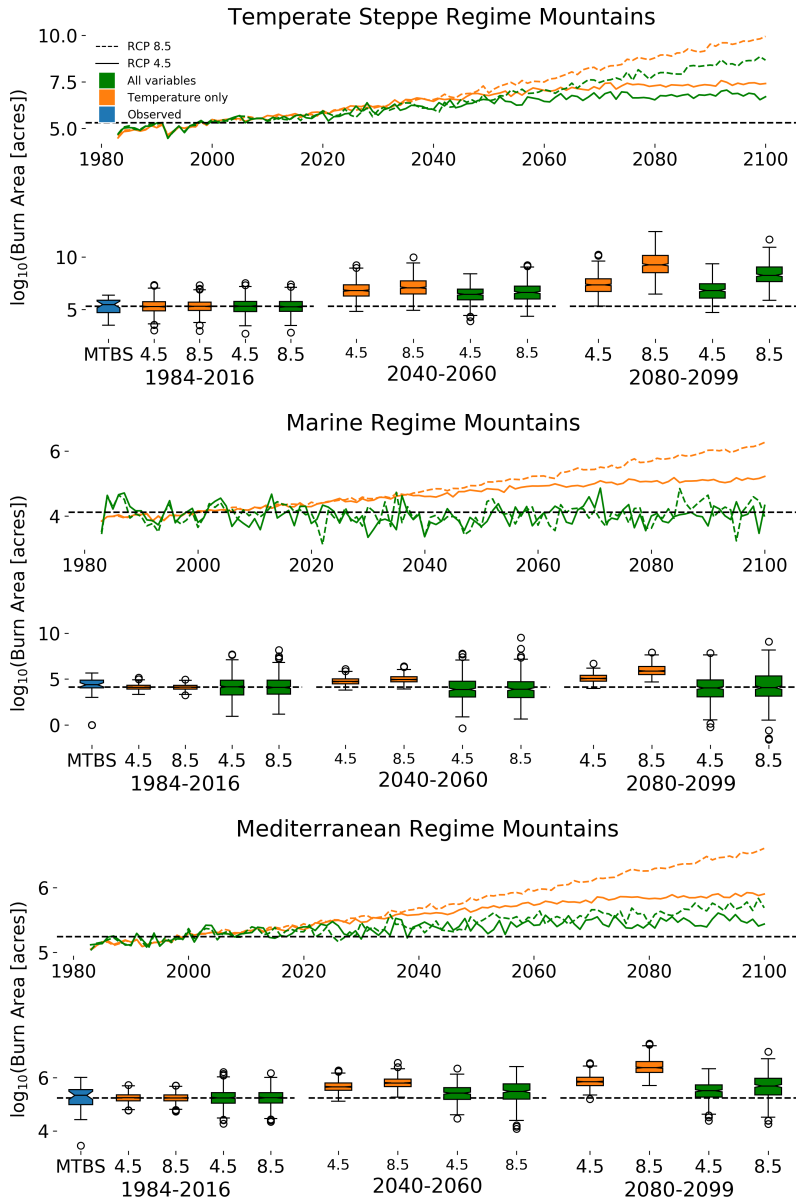
FIG. 3.6. Temperate Steppe Regime Mountains (top), Marine Regime Mountains (middle), and Mediterranean Regime Mountains (bottom) predicted burn area based on Lasso regressions trained with all variables (Lasso-all, green) versus just temperature (Lasso-Temperature, orange) for RCPs 4.5 and 8.5. The line plots show the ensemble mean of individual CMIP5 model burn area projections for a given regression type and RCP. We choose not to show historical burn area (blue line) on this figure, as comparing an ensemble mean (plotted curves) against a single realization (observed burn area) is not straightforward, and prone to confusion, given that there are more wobbles in a single realization than an ensemble mean. Below the time series, box plots show the distribution of individual summer burn area estimates from all available CMIP5 models for the years indicated on the horizontal axis label. These distributions are separated by regression type (color) as well as RCP (horizontal axis tick labels).

Due to the large and nearly certain increases in temperature observed across CMIP5 models and RCPs, regularized regression using only seasonal temperature values predictors produce much larger future burn area increases than the Lasso-All regression, where temperature is not always chosen. The variability of these estimates is also low. The difference between the two regressions is smallest in the Temperate Steppe Regime Mountains, where temperature was objectively identified as a predictor for the Lasso-All regression. No temperature variables were selected in the Marine Regime Mountains and without temperature, the burn area estimates have large variance, no clear trend, and very little difference between RCPs. The boxplots of individual summer burn area estimated for future decades in the Marine Regime Mountains suggest future summers could have near-zero burn area, or burn area totals 2 orders of magnitude larger than any summer in the historical period, depending on what combination of conditions occur. In the Mediterranean Regime Mountains, winter temperature was selected as an important feature and is responsible for most of the increase in burn area seen for the Lasso-All regression (shown in Figure A11). Variables that depend on temperature but explicitly track moisture in the environment (evaporation, soil moisture, precipitation, relative humidity) have much more uncertain future mean states than temperature (Figure 3.5). This is clearly demonstrated in the Lasso-All projection for the Marine Regime Mountains, where temperature was not selected. During the historical period, warm summers were associated with more burn area than cool summers. This relationship is captured by the Lasso-Temperature regression, but not always essential for predictive power when other wildfire relevant variables are available.

As discussed previously, the best fit regressions tend to underestimate the largest and overestimate the smallest summer burn area totals. Even if our regressions had R^2 values equal to 1, the spread displayed by the boxplots in Figure 3.6 would be an underestimate, because our limited data, covering only 1984 through 2016, do not sample the full range of possible values (phase space) for environmental conditions or summer burn area. CMIP5 models offer a more comprehensive sample of the environmental phase space. Combined, the CMIP5 models provide many more years of what these data could look like. This is how the CMIP5 driven burn area distributions can be wider than the observations distribution while still underestimating total spread. This is shown by the boxplots in Figure 3.6, which show the distribution of summer burn area for MTBS and CMIP5 driven regression estimates for the years 1984-2016. The CMIP5 burn area distributions can also be narrower than the observations. This is most likely when the regression explains a small proportion of the historical variance. Thus,

the spread for CMIP5 driven summer burn area for all regressions for the decades 2040-2060 and 2080-2099 shown in Figure 3.6 are more narrow than would be the case if our historical data covered the full phase space.

The burn area projections and distributions shown in Figure 3.6 are meant to highlight how estimates of future wildfire burn area are sensitive to the environmental conditions, RCPs, and global climate models used to make the estimate. Though exponential relationships between burn area and changing environmental conditions have been observed in the recent past, extrapolating these relationships into the future almost certainty overestimates future wildfire burn area. In the past, western U.S. forest wildfire activity has been flammability limited (Abatzoglou and Williams 2016), however, ever increasing burn area would eventually result in limited fuel availability. The limited observational data record available for the regression analysis does not capture the true scope of the wildfire-climate relationship. Even if it did, the observed relationships are unlikely to remain unchanged in the future. There is evidence that in recent decades changes in land management and the legacy of wildfire suppression resulted in heightened responses of wildfire burn area to variability in environmental conditions that may not persist into the future (Higuera et al. 2015; Littell et al. 2016). Finally, the relationships presented here are representative of large spatial and temporal scales. The relationships shown here may not apply to finer scales or individual wildfires. At smaller spatial and temporal scales, the impact of fire-weather (windy days with low RH%) on wildfire occurrence and spread is reasonably well understood, but these scales are not resolved by climate models. Future large-wildfire occurrence is still subject to the co-occurrence of an ignition source, dry fuels, the absence of substantial rainfall, and high winds, a set of conditions this work does not explicitly estimate.

This work shows that the estimates for how burn area may change in the future is extremely sensitive to what environmental predictors are chosen to drive burn area. We show that increases in burn area are likely, as changes in temperature, as well as other candidate variables, indicate the future will likely be drier. CMIP5 models are in agreement that temperature will increase in the 21st century. Temperature is the only variable that increases across all ecoregions, models, and RCPs, thus using temperature as a key predictor of future burn area will likely underestimate uncertainty. Future wildfire burn area estimates are lower whenever and wherever the importance of temperature is reduced.

3.4 DISCUSSION

There are limitations to this work, and thus trends in future western U.S. summer burn area could be significantly different from patterns presented above. Burn area interannual variability may not scale with climate change the way it has in the historical period. The climate models and regional scales used in this work are only able to capture changes at large scales; however, there are subsets within western U.S. ecoregions that may in fact cool in the future as there is a decoupling of scales due to topography (Daly et al. 2009). This work uses a single CMIP5 ensemble member (r1i1p1) for each CMIP5 model, which limits the internal climate variability probed by this analysis. This work only examined changes in summer wildfire burn area. We did not consider how the length of the wildfire season is likely to increase (Yue et al. 2013b), or how that could impact future burn area. Our work shows that changes to environmental variables in the RCP 8.5 scenario result in larger wildfire potential than RCP 4.5. Some work has demonstrated that RCP 4.5 could lead to episodically larger wildfires in the future as that pathway has conditions conducive to fuel build up (Bachelet et al. 2016). Our methods do not account for feedbacks like this. Much of the historical emphasis on temperature has come from its impact of vapor pressure deficit (VPD), since this is a good proxy for moisture demand from dead fuels, and many wildfires are carried by fine dead fuels (fuels $\frac{1}{4}$ inch in diameter or less) (Anderson and Rothermel 1965), though this does not account for extreme wildfire behavior such as crowning. Thus, it may not matter how water stressed live fuels may or may not be, VPD, which scales exponentially with temperature, could accurately account for flammability.

An additional area of weakness for this and other studies that leverage statistical models is their reliance on stationary relationships to make predictions. Large-scale forest- and land-change disturbances are already happening in the west (Murph and Mooney 2019), the fuel available to burn, as well as wildfire-climate dynamics are likely to change. In some experiments, vegetation models show large shifts in vegetation towards warmer types (*e.g.*, temperate to subtropical forest types, warm subtropical grasslands replacing cool temperate grasslands) as the west warms (Harris et al. 2016). Wildfires themselves could become a source of disturbance that could shift species towards those that are more well adapted to the new climate. It is also possible that fuel density could reduce if trees begin to die off due to temperature stress (Park Williams et al. 2012). Decreases in forest productivity have been observed globally and attributed to water limitation (Allen et al. 2010; Zhao and Running 2010). However, there is also evidence that plant physiology responses to increased CO_2 concentrations may reduce live fuel stress and flammability. Plants absorb CO_2 through stomata in their leaves and lose water to

the atmosphere through the same pathway (Swann et al. 2016). Some plants lose less water per unit of carbon gain when CO_2 mixing ratios increase because the gradient of CO_2 between the leaf and atmosphere is reduced (Cowan 1978). However, plant water-use efficiency will only increase if the loss of water per unit of carbon gained is not offset by an increase in leaf area (Field et al. 1995). Assuming nearly constant leaf area, plant increased water use efficiency could result in reduced transpiration and increased soil moisture (Field et al. 1995). Combined, these changes could reduce plant water stress, even during droughts. Some observations show a decrease in transpiration due to increased water use efficiency (WUE) with increasing CO_2 mixing ratios (Keenan et al. 2013; Peñuelas et al. 2011; van der Sleen et al. 2014; Warren et al. 2011), though the number of plant species this applies to and limits are not fully understood (Battipaglia et al. 2013).

Despite these shortcomings, this work offers novel contributions. 1) We objectively select variables that explain historical year-to-year variability in the western U.S. burn area, allowing the data to speak for itself rather than confirming an existing hypothesis about one variable being the most important. This approach shows that temperature is not uniquely suited to give statistical models the most skill. 2) We show how historical relationships between variables and changes to those variables simulated by CMIP5 may contribute to future burn area for multiple western U.S. ecoregions and RCPs. 3) We demonstrate that future estimates of wildfire burn area are sensitive to what environmental variables are chosen to predict it. 4) We show that smaller increases in future wildfire burn area are estimated whenever and wherever the importance of temperature as a predictor is reduced.

CHAPTER 4

ESTIMATING THE SPREAD IN FUTURE FINE DUST CONCENTRATIONS IN THE SOUTHWEST UNITED STATES

4.1 INTRODUCTION

Airborne dust can be a major component of western United States (U.S.) particulate matter (PM) pollution, with larger dust particles impacting PM_{10} and fine dust influencing $PM_{2.5}$ (particles with diameters smaller than 10 and 2.5 μm respectively) concentrations. Dust aerosol can impact the Earth's radiation budget, heterogeneous chemistry, cloud formation, and human health (Krueger et al. 2004; Grineski et al. 2011; Miller and Tegen 1998; Rosenfeld and Nirel 1996; DeMott et al. 2003; Bond et al. 2013). U.S. natural dust emissions (wind blown dust of natural non-agriculture lands) are the largest in the western U.S., where large deserts and bare ground are located. In the southwest U.S. (an area encompassing Utah, Colorado, Arizona, and New Mexico), local sources of dust include the Mojave, Great Basin, Sonoran, Chihuahuan deserts, and the Southern Great Plains. Dust emissions from these regions are responsible for the highest wind-blown dust concentrations in the U.S. (Prospero 2002). Wind blown desert dust is comprised of mostly silica, which is known to cause chronic lung inflammation, fibrosis, and lung cancer (Sing and Sing 2010; Prüss-Ustün et al. 2011; Steenland and Ward 2014). Achakulwisut et al. (2018) estimated that southwest U.S. fine dust concentrations will increase 0.15 - 0.58 μg m^{-3} under representative concentration pathway (RCP) 8.5, resulting in approximately 750 premature deaths annually. Dust has impacts even after it leaves the atmosphere. When snow is covered by dust, it absorbs more solar radiation and melts faster because the shortwave absorption coefficients of dust are much larger than snow (Warren and Brandt 2008); as a result, dust has been shown to play a more important role than temperature in the timing of snowmelt in the Rocky Mountains (Painter et al. 2018). These changes in snowmelt can influence the timing of water availability in the Colorado river (Painter et al. 2018). The timing of spring snowmelt in the western U.S. has been linked to increased wildfire activity (Westerling 2016; Westerling et al. 2006). Dust also impacts visibility in the southwest (Achakulwisut et al. 2018; Hand et al. 2016).

There are many environmental variables that have been shown to influence dust emissions and concentrations in the southwestern U.S. and elsewhere. The presence of vegetation can greatly reduce soil erosion by wind and can limit dust emissions (Marticorena and Bergametti 1995; Woodward et al. 2005; Pu and Ginoux 2017; Kim et al. 2017). Vegetation cover is an important control for dust emissions and is influenced by precipitation, however, vegetation cover is slow to respond to precipitation (Evan

et al. 2016). Wet (heavy) soil is harder to loft than dry soil, making soil moisture a controlling factor of dust emissions (Fécan et al. 1998). Precipitation can increase soil moisture and scavenge airborne dust (Ginoux et al. 2001).

There are also demonstrated links between dust and the El Nino-Southern Oscillation (ENSO), and the Pacific Decadal Oscillation (PDO), due to their impact on precipitation and temperature, and transpacific transport (Achakulwisut et al. 2017, 2018). Pu and Ginoux (2017) show that variations in precipitation, soil bareness, and surface wind speed control dust emissions. The importance of surface wind and plant cover are also demonstrated by Kim et al. (2017). Evan et al. (2016) show that monthly dust emissions are linearly related to mean monthly wind speed. However, while observationally based analyses point out key variables that influence dust emissions, process-based models struggle to reproduce observed dust emissions and concentration variability (Mahowald et al. 2010; Pu and Ginoux 2018). Models in Phase 5 of the Coupled Model Intercomparison Project (CMIP5) are able to capture global mean dust optical depth (Pu and Ginoux 2017). However, not all CMIP5 models use an interactive dust scheme (Pu and Ginoux 2017), making the use of statistical models necessary to increase the number of models available to estimate intermodel spread. Pu and Ginoux (2017) found that the CMIP5 models that do have interactive dust schemes tend to estimate a more dusty future than regression analysis, while failing to capture the observed relationships between dust optical depth with surface wind speed, bareness, and precipitation.

Due to high dust concentrations compared to the rest of the U.S., a long record of fine dust observations, and anticipated population growth, this work focuses on fine dust in the southwestern U.S. We aim to understand how estimates of future dustiness depend on the choice of environmental predictors used to make the estimate. The southwest is not the only U.S. region with large dust emissions, however, other large dust source regions such as the Southern Great Plains are mostly anthropogenic in origin (Lee et al. 2009, 2012; Bullard et al. 2011) and our work is focused on the biophysical controls of natural dust emissions. Our study region covers Utah, Colorado, Arinoza, and New Mexico (see Figure 4.1). A regional approach is appropriate for this work because fine dust tends to have regional impacts relative to the more localized impact of coarse dust (Hand et al. 2016). Further, this work is motivated by the estimated health impacts of seasonal fine dust concentrations (*e.g.* Tong et al. (2017); Achakulwisut et al. (2019)). Achakulwisut et al. (2018) use seasonal fine dust concentrations as a metric for dust burden, and estimate health impacts. This work is motivated by these findings and aims to build upon the discussion and exploration of uncertainty of the environmental conditions that drive seasonal dust

concentrations. Further, our work aims to understand long-term changes in dust drivers with a large ensemble of CMIP5 models. We can more easily achieve this goal working with monthly data than daily data, and these will be more closely related to seasonal $PM_{2.5}$ concentrations than episodic events (*e.g.* dust storms).

The specific goal of this work is to understand how estimates of future dustiness depend on the choice of environmental predictors used to make the estimate. By doing so, we explore and estimate a range of possible future dust concentrations. We use CMIP5 models to estimate an intermodel spread of dust drivers and their influence on future dustiness for Representative Concentration Pathways (RCPs) 4.5 and 8.5, scenarios that correspond to different greenhouse gas climate forcing ($W\,m^{-2}$) scenarios for the 21st century. We use environmental variables identified by other studies (hereafter referred to as "candidate variables") that drive southwestern U.S. dust emissions and concentrations. This work measures dustiness based on concentrations rather than emissions due to the availability of high quality dust concentration measurements and the air quality motivation for this work. Our approach is designed to comprehensively quantify spread in both the drivers of and concentrations of future dust. We address the following questions: 1) How do candidate variables change in the future as estimated by CMIP5 models for RCP 4.5 and 8.5? How does intermodel spread compare to scenario (RCP) spread? 2) What changes to future dust concentrations do historical linear relationships between dust and key environmental variables predict, and how do these estimates and their spread compare between variables? 3) How does a linear combination of candidate variables, objectively subset to include only those that maximize predictive skill on unseen data, compare to the single variable estimates of future dust season concentrations?

4.2   DATA AND METHODS

4.2.1   *Quantifying dust using IMPROVE measurements of Fe*

We quantify fine dust concentrations using the method put forward by Hand et al. (2016), which assumes fine dust to be 3.5% iron (Fe) based on crustal abundances (Taylor 1985). Hyslop et al. (2015) demonstrate that Fe concentration measurements have been performed consistently over time. In the Southwestern U.S., Fe emissions are primarily a result of mineral sources (Wang et al. 2015). Fe concentrations are measured by monitors in the IMPROVE (Interagency Monitoring of PROtected Visual Environments) network (Malm et al. 1994), which collect aerosol samples for 24 hours every third day

at predominantly rural and remote sites. Mass concentrations are recorded as mass per air volume at local temperature and pressure.

We use a subset of IMPROVE sites available in the southwest region based on data availability. We use monitoring sites that have nearly continuous measurements suitable for monthly mean estimates available in the southwest region between 1995 and 2015. IMPROVE data are available in more recent years, but we subset the data to match the data range of the Leaf Area Index (LAI) data used in this analysis (described next). These criteria are met by 15 monitoring site locations (site details in Appendix A12). We use an additional 2 monitors that have gaps in their data record in order to provide additional geographic coverage in New Mexico and Northwest Arizona. On average, monitors located further south have higher monthly mean concentrations (Appenix A13). We do not attempt to distinguish between anthropogenic and natural dust sources, or the possible influence of long range transport. IMPROVE data were downloaded from: http://vista.cira.colostate.edu/improve/Data/.
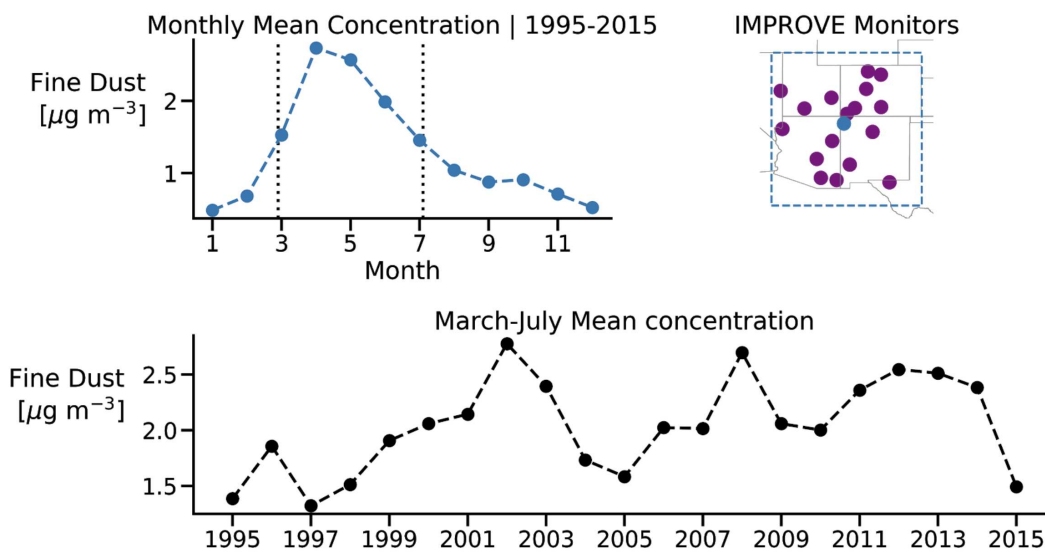

FIG. 4.1. Summary of the dust data used in this analysis. Topleft: Monthly mean dust concentrations measured by IMPROVE monitors in the southwest U.S. between 1995 and 2015. The dashed vertical lines show the "dust season" that we focus on for this work, which include March through July. Topright: Locations of IMPROVE monitors used in this work (purple) and the geographic centroid of the monitors (blue), as well as the study region (dashed blue border). Bottom: Southwest dust season (March - July) mean fine dust concentration ($\mu g\,m^{-3}$) measured by the IMPROVE monitors plotted in the upper right panel.

Figure 4.1 shows a summary of the data used to estimate southwest (Figure 4.1, top right panel) fine dust seasonal concentrations in this analysis. We focus on months March through July (hereafter "dust

season"), as these months have the highest dust concentrations (Figure 4.1, top left panel). The bottom panel shows the time series for the mean dust season concentration as measured by the monitors shown in the upper right panel of Figure 4.1.

4.2.2   Leaf Area Index (LAI)

Leaf Area Index (LAI) is a dimensionless metric that expresses the ratio of one sided leaf area per unit area of ground (m^2/m^2). LAI influences land-surface conditions, exchanges of water, energy, and CO_2 between land/vegetation and the atmosphere (Brovkin 2002; Mahowald et al. 2016). In North America, LAI values typically range between 0 (bare ground) and 10 (dense forests). LAI has been shown to be anticorrelated with dust concentrations in the Southwest U.S. and elsewhere (*e.g.* Pu and Ginoux (2017); Hand et al. (2016); Kim et al. (2017)). We use LAI estimates derived from version 4 of the Climate Data Record (CDR), which are based on surface reflectance observations from the National Oceanographic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) (Claverie, Martin, Eric Vermote, NOAA CDR Program (2014) 2014). These observations are made from seven NOAA polar orbiting satellites. The data are stored on a grid with 0.5 degree horizontal grid spacing. Data availability varies by location and month, which results in monthly mean LAI values for the southwest region being composed of differing numbers of daily observations. The data are available from 1981 to the present; however, at the time of this analysis, 2016 LAI data were only available as "preliminary data", so we only use data through 2015. These data are available at the following URL: https://www.ncei.noaa.gov/data/avhrr-land-leaf-area-index-and-fapar/access/.

We use CMIP5 model output for future LAI estimates. We find that CMIP5 models simulate larger monthly mean LAI values than are observed by the AVHRR product, and this has been documented by others. Anav et al. (2013) show that climate models (coupled and non-coupled) overestimate seasonal average LAI values. We bias correct the mean and year-to-year variance of CMIP5 estimates of LAI to match the satellite-based observations (described in greater detail below). Several CMIP5 models show an upward trend in LAI during the historical time period that differs from observed trends.

4.2.3   ECMWF ERA-Interim Reanalysis

For historical estimates of all candidate variables aside from Leaf Area Index, we use the European Center for Medium-Range Weather Forecasts (ECMWF) ERA-Interim monthly reanalysis fields Dee et al. (2011). We use temperature, total precipitation, wind speed, relative humidity (calculated using dew point temperature, surface pressure, and Tetens formula), and soil moisture. All data were

downloaded at the native horizontal grid spacing of 0.75° x 0.75°. To make the ERA-Interim output directly comparable to CMIP5 output (described next), variables were bilinearly re-gridded to a common grid with 2° x 2.5° degree grid spacing (Schulzweida 2019). Where necessary, accumulation units ($e.g.$ m day^{-1}) were converted to flux units ($e.g.$ kg m^{-2} s^{-1}) to match the units used by CMIP5 models. A total root zone (3 meter depth) soil moisture field was created by converting volumetric soil water in layers 1 through 4 (total depth of 2.89 m) to kg m^{-2} by assuming the density of all water content to be 1000 kg m^{-3}.

4.2.4   CMIP5

To estimate how candidate variables change in the future, we use output from the Phase 5 of the Coupled Model Intercomparison Project (CMIP5) (Taylor et al. 2012). We use the following monthly variables: near surface temperature, precipitation, near-surface wind speed, near-surface relative humidity, soil moisture content, and LAI. Intermodel spread between models is quantified using the r1i1p1 ensemble member for each model. We use a single ensemble as our focus is quantifying scenario and intermodel spread. Scenario spread is quantified using output from representative concentration pathways (RCP) 4.5 and 8.5, scenarios that correspond to different greenhouse gas climate forcing (W m^{-2}) scenarios for the 21st century. RCP 4.5 represents an emission scenario where greenhouse gas emissions begin to slow in the mid 21st century while RCP 8.5 represents a scenario where emissions continue to grow throughout. In order to bias-correct CMIP5 data using the reanalysis data, historical (pre-2006) CMIP5 data were also used. The data from all models were regridded to the GFDL-CM3 2° x 2.5° degree grid using bilinear-interpolation for atmospheric domain variables and distance-weighted average remapping for land domain variables (Schulzweida 2019). Data are averaged over the southwest region plotted in Figure 4.1. Data for this work was downloaded from the Earth System Grid Federation (https://esgf-node.llnl.gov/search/cmip5/). We vertically interpolate and integrate soil moisture down to the same depth as ERA-Interim.

4.2.5   Linear relationships between candidate variables and dust concentrations

We use the Least Absolute Shrinkage and Selection Operator (hereafter Lasso) regression to create regularized linear models to estimate linear relationships between candidate variables and average southwest fine dust concentrations in the months March through July. The Lasso is a regression method that performs regularization and variable selection (Tibshirani 1996). It was developed to increase the prediction accuracy of ordinary least squares (OLS) regression by reducing the variance of predicted values (Tibshirani 1996). This reduction of the variance of predicted values is accomplished

through a constraint on the sum of the absolute values of the model parameters, which results in some predictor coefficients being set to exactly zero. The Lasso has been demonstrated to make a good compromise between model simplicity and model performance (Loukina et al. 2015). The optimal constraint on the sum of model parameters is selected using leave-one-out cross validation, where the model with the lowest mean square error on left-out data is used for predictions. We implement these methods using the LassoCV method from scikit-learn Pedregosa et al. (2011). All candidate variables are standardized (mean 0, unit variance) before regression analysis.

For the multilinear regression component of this work, we implement the so-called "relaxed Lasso", where Lasso regression is used twice (Meinshausen 2007). This relaxed Lasso is done in two steps. First, we use Lasso regression for variable selection. This initial set of regressions is done by leaving out a year of data, fitting a Lasso regression using leave-one-out cross validation on the remaining years and recording what variables receive a non-zero coefficient. We repeat this process for each year. We discard the 50% of variables that were selected (non-zero coefficient) the least often. This step helps us identify what proportion of the time a given variable is selected, which gives a measure of how robust that variable is to noise in the yearly data. Second, we fit the Lasso regression using leave-one-out cross validation using the 50% of the original features that we selected most frequently in step 1.

In our use of Lasso regression, environmental variables from ERA-Interim reanalysis and satellite observations are used as predictor features. These features include all of the candidate environmental variables listed in the introduction and others related to aridity (relative humidity and vapor pressure deficit), where the average value for each feature is taken for the dust season (March through July). ERA-Interim data are spatially subset to include only the grid boxes that overlap the southwest region, shown in Figure 4.1. Due to the limited number of years with all required data (1995 to 2015), we set up our regression analysis to explain variance in year-to-year dust concentrations. One of the underlying assumptions of this approach is that year-to-year variability relationships can be leveraged to estimate long-term changes in dust season concentrations. To avoid a spurious model where the target function and predictors are well correlated because of a shared trend, we detrend the candidate variables and the dust concentration time series before fitting the regressions. This ensures that the multi-linear regression predicts year-to-year variability and that correlations between concentrations and candidate variables are due to interannual co-variability.

We use CMIP5 model output and our Lasso regressions to estimate how dust concentrations may change in the future. We bias-correct CMIP5 model output such that the year-to-year standard deviation and mean value for the historical period (1995-2015) match those estimated by the ERA-Interim reanalysis and satellite derived LAI estimates. This was done for each candidate variable predicted by each CMIP5 model individually using the following procedure. 1) The linear trend is removed from the 1995-2015 CMIP5 variable time series (*e.g.* summer precipitation for the ACCESS1-0 model). 2) The standard deviation, which after detrending is due entirely to year-to-year variability, is calculated for the detrended 1995-2015 time series. 3) The detrended 1995-2100 CMIP5 time series is multiplied by the ratio of the detrended ERA-Interim variable standard deviation over the detrended CMIP5 variable 1995-2015 standard deviation. At the completion of this step, the CMIP5 variable standard deviation for 1995-2015 matches the historical estimate. 4) The 1995-2100 linear trend is added back to the CMIP5 feature. 5) To correct the CMIP5 model bias offset, the mean value of the time series from 1995-2015 is replaced by the ERA-Interim mean for the same years. We choose to bias correct CMIP5 output in this way because it eases the interpretation of how variables are estimated to change in the following ways. 1) It allows for a straightforward interpretation of changes observed by models from the historical period to the simulated future. 2) It gives us the ability to observe how the spread in a variable changes due to changes in the range of year-to-year values. 3) It ensures the spread in the estimates of future dust concentrations estimated by CMIP5 output using regression models is not artificially narrow due to a small variance in CMIP5 model variables.

4.3 RESULTS

4.3.1 *Simulated changes in candidate variables and future dustiness*

As discussed in the introduction, there are many environmental conditions that have been shown to influence dust emissions and concentrations. This section shows how each candidate variable is predicted to change from the historical period (1995-2015) to the end of the 21st century (2070-2099) as estimated by CMIP5 model output for RCP 4.5 and 8.5. This change is estimated for each model and RCP by taking the mean candidate variable values for all dust seasons (March through July) in the future period minus the mean value for the historical period. Due to consistent dust data availability in the southwest, the historical period starts in 1995, and goes through 2015, the last year of LAI data available. In addition, we estimate the linear relationship between candidate variables and dust season concentrations for the years 1995 through 2015. Linear relationships are established for each candidate

variable and dust season concentrations using Lasso regression. Only a single predictor is used in this section, resulting in only a small amount of regularization, so the linear relationships presented in this section are nearly identical to those that would be estimated by ordinary least squares regression. A summary of these regressions variance explained and Pearson correlation coefficients are shown in table 4.1.

TABLE 4.1. The historical variance explained (R^2) and pearson correlation coefficient (r) of the Lasso regression for single variable relationships between historical data and observed seasonal fine dust concentrations. Correlation coefficient shows the sign of the correlation.

$Variable$	R^2	r
Temperature	0.147	0.383
Precipitation	0.450	-0.673
Wind Speed	0.107	0.327
RH%	0.486	-0.700
VPD	0.355	0.596
Soil Moisture	0.340	-0.585
Leaf Area Index	0.213	-0.462

Table 4.1 and Figures 4.2 and 4.3 present a comprehensive view of how candidate variables are related to dust in the historical period, simulated to change from the historical period to the future period (Figure 4.2), and what those changes imply for future dust season concentrations based on the historical linear relationships between environmental variables and dust concentrations (Figure 4.3). Simulated changes for variables for RCP 4.5 and 8.5 mostly agree on the sign of the change from the historical to future period, though changes for RCP 8.5 are usually larger than RCP 4.5 (Figure 4.2). Temperature and vapor pressure deficit (VPD) are the only variables where every model for both RCP scenarios simulate increased values. Both are positively correlated with seasonal dust concentrations. Leaf Area Index (LAI), a quantity anticorrelated with bareness and dust season concentrations, is simulated by the fewest models, but has the largest intermodel spread of all candidate variables. The large spread in LAI simulated across models and RCPs is consistent with other studies (*e.g.* Mahowald et al. (2016)). The changes in LAI shown in Figure 4.2 have been multiplied by a factor of 1/3 in order to present changes in other variables on a scale that clearly presents model spread. Only three models estimate decreased LAI at the end of the 21st century, two for RCP 4.5 and one for RCP 8.5.
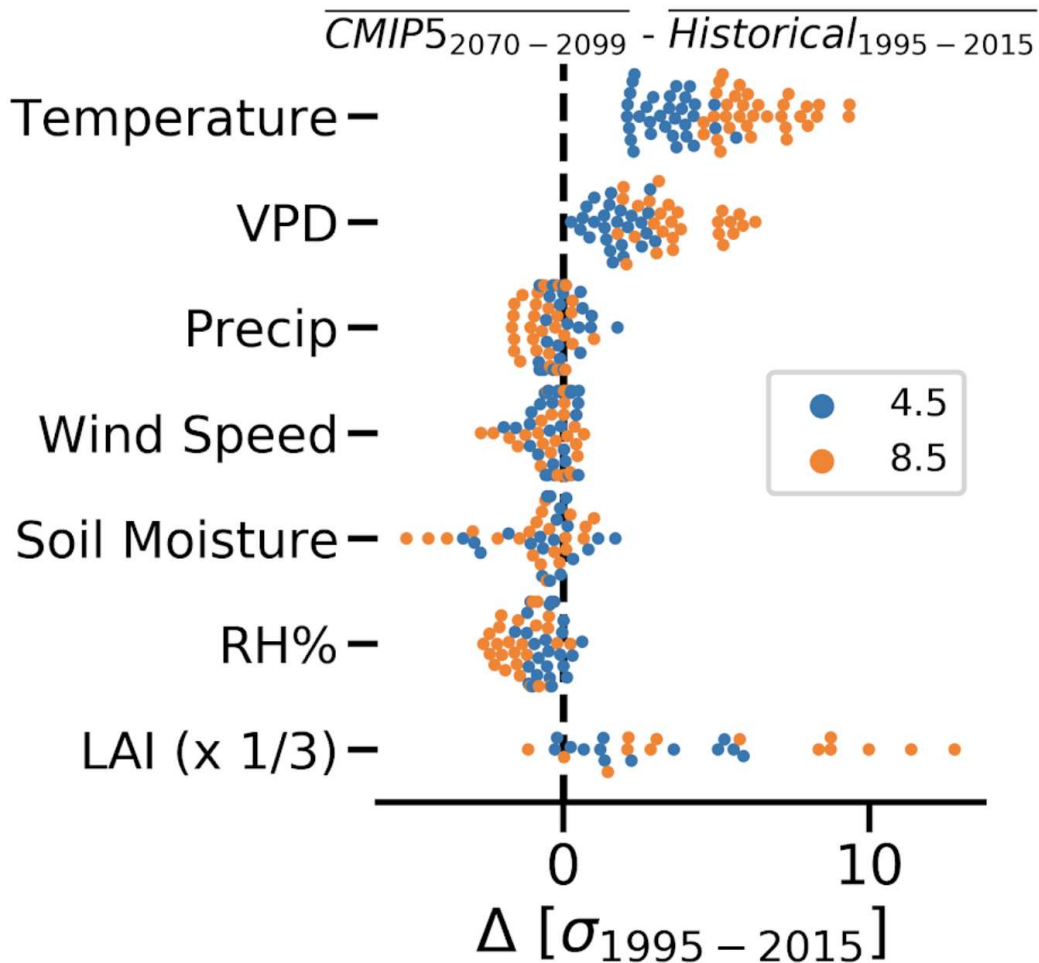
FIG. 4.2. The change in model estimated mean dust-season values from 1995-2015 to 2070-2099. Each variable is expressed in units of the detrended historical period (1995-2015) year to year standard deviation (σ). Colors correspond to RCP 4.5 (blue) and 8.5 (orange). Points that fall on the dashed vertical black line indicate models that simulated no change for the mean of variables for the dust season between the historical and future period. Changes in Leaf Area Index (LAI) have been multiplied by 1/3 in order to reduce the size of the horizontal axis.

Precipitation and soil moisture are anticorrelated with dust emissions and concentrations (Table 4.1). Though some models show increases, most simulate less precipitation and soil moisture by the end of the 21st century, but this drying is particularly pronounced for RCP 8.5. The soil moisture in this work goes to a depth of nearly three meters, a depth corresponding to the so-called root-zone soil moisture, which as a whole is less sensitive to surface conditions (*e.g.* VPD) than soil moisture in the first few centimeters, as water at depth can generally only be removed through plants. The majority of CMIP5 models show decreased future soil moisture for both RCP scenarios. Wind is the mechanism

that lofts dust from the surface and the regression shows this relationship with wind being positively correlated with dust emissions. The majority of CMIP5 models simulate decreased dust-season wind speed in the future. Of the candidate variables shown in this work, most relate to dust through their influence on the dryness and loftability of dust. No matter what the surface conditions are, dust is only lofted when it is windy, and most models show a reduction in wind speed in the future.
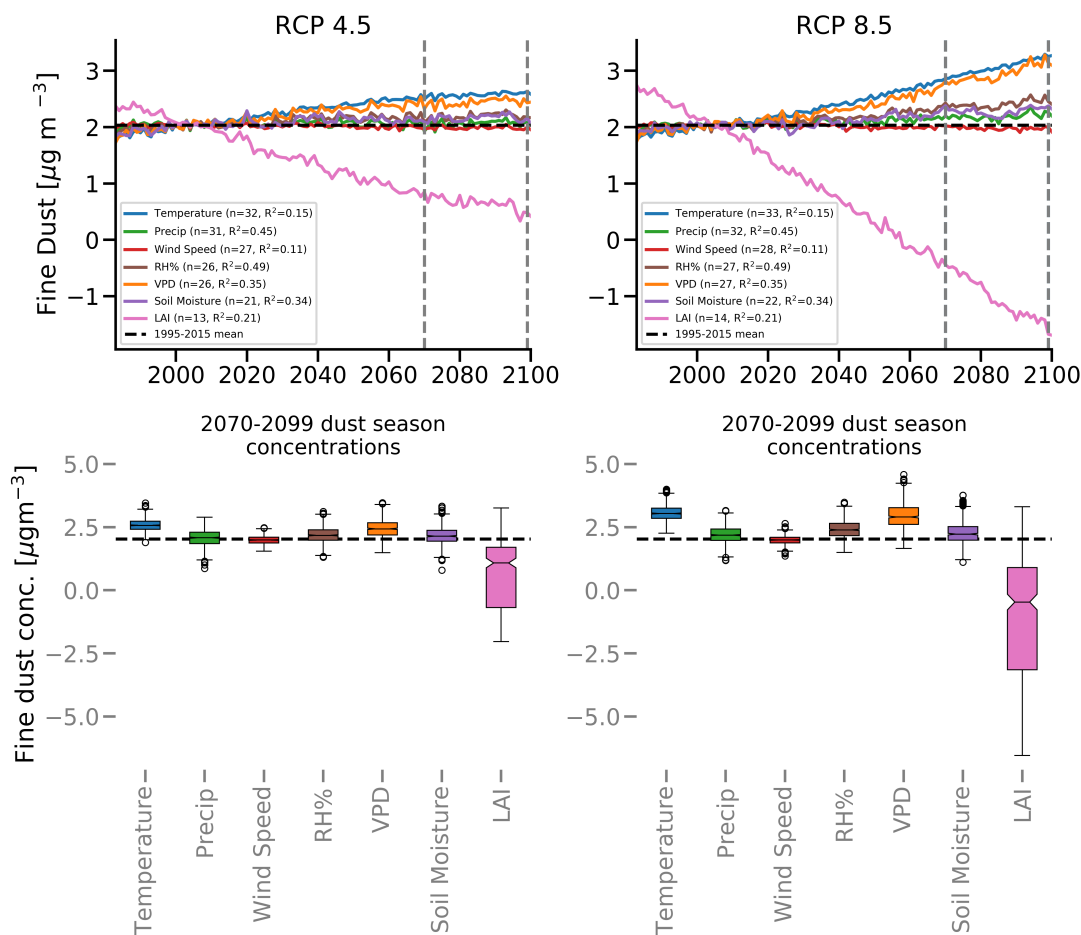

FIG. 4.3. Top row, future dust season (March-July) concentration estimated by single variable Lasso regression for RCP 4.5 (left) and RCP 8.5 (right). The lines show the ensemble means of all available models (n) for a given environmental variable. The bottom row shows the distribution of individual model estimates for individual dust season mean dust concentration in the years 2070-2099 (years indicated in times series with vertical dashed lines for each RCP). For the top and bottom, the dashed black horizontal line represents the 1995-2015 southwest dust mean concentration. The dust season concentration distribution shown in the bottom panel correspond to the CMIP5 changes shown in Figure 4.2.

Figure 4.3 shows the CMIP5 model ensemble mean dust concentrations for each candidate variable predicted through our analysis. Estimates of future dust concentration estimates are driven by the

linear relationships between each individual variable and dust concentrations in the historical period. None of the linear relationships between candidate variables and dustiness explain more than half of the historical variance observed in the historical period. Relative humidity explains the most historical variance (R^2=0.49) and wind speed explains the least (R^2=0.11). However, despite the low proportion of variance explained, we can determine the sign of the relationship between the candidate environmental variables and determine what changes these relationships imply for future dust concentrations as well as how they compare to each other (Table 4.1). For example, the multimodal mean change in temperature and VPD show the largest increase in future dust concentrations for both RCPs. We know from Figure 4.2, that every model for both RCP scenarios simulates increases in these variables.

The simulated CMIP5 LAI-driven changes in future dust concentrations show that increasing LAI (Figure 4.2) results in decreased dust season concentrations (Figure 4.3). We have low confidence in the CMIP5 LAI-driven projections of dust concentrations for two reasons. 1) LAI is not well correlated with mean dust season concentrations. 2) The future estimates of LAI values have higher intermodel spread than any other candidate variable. Though physically important for whether dust emissions can occur (as indicated by other work) the spread presented by LAI suggests that if it is the critical control on dust in this region, future dust concentrations are likely to decrease. Mahowald et al. (2016) show that CMIP5 coupled ESMs that simulate LAI values that most closely resemble satellite observations also tend to be the models that simulate smaller increases in the future, though these results were for the tropics. By contrast, relative humidity, which is well correlated with temperature, explains roughly half of the historical year-to-year variance of dust concentrations and most CMIP5 models simulate decreased relative humidity in the future for both RCP 4.5 and 8.5 (Figure 4.2), which based on historical relationships, imply increased future dust concentrations (Figure 4.3). The spread in dust concentrations shown by the boxplots shown in Figure 4.3 are an underrepresentation of future spread for two reasons. First, due to the limited temporal range of the observational data, the historical data are unlikely to account for the true variability or range of values for dust season concentrations, or the corresponding state of candidate variables (*i.e.*, observations do not sample the full phase space). Second, even if the historical period did sample the full phase space, none of the single variable regressions explained all of the observed historical variability (R^2<1).

4.3.2 Relaxed Lasso estimates of future dustiness

Southwest dust emissions and concentrations are influenced by more than a single environmental variable in isolation. In this section, we use a custom implementation of relaxed-Lasso linear regression (described in the data and methods section) to objectively select a linear combination of candidate variables and estimate future dust-season concentrations. The linear combination of variables selected by the relaxed-Lasso regression that minimize square error on unseen data are precipitation, wind speed, relative humidity, and soil moisture. The linear model explains 66% of the observed variability in mean dust concentrations for the historical period, more than any single variable in the previous section (max $R^2 = 0.49$). The regression coefficient for each selected variable is shown on the horizontal axis of Figure 4.4 as well as Table 4.1 (additional regression details are shown in Figures A14 and A15). Expressed in terms of year-to-year standard deviation, changes in precipitation and relative humidity most strongly influence dust concentrations. Figure 4.4 shows the relaxed-Lasso regression coefficients against the CMIP5 multimodel mean estimate of how dust season conditions are simulated to change from the historical period (1995-2015) to the end of the 21st century (2070-2099). The dots indicate the mean change across individual models and the whiskers show the total range of change estimated across all of the models for each RCP scenario. Dots in the red shaded areas indicate that in the context of the relaxed-Lasso regression, the multi-model mean change in that variable will result in increased dust concentrations, while dots in the white areas will result in decreased dust concentrations. Between RCP 4.5 and 8.5, most of the selected variables change in a way that the regression indicates will lead to higher dust concentrations in the future, though all variables show at least one model that has the opposite impact (whiskers). On the other hand, reductions in dust season wind speed works to reduce dust concentrations for both scenarios, but some models show the opposite impact. The multimodel mean change in RCP 4.5 precipitation shows a slight increase from the historical period, but the mean change for RCP 8.5 shows a larger decrease in future precipitation. The intermodel spread in estimates of how variables will change is largest for soil moisture.
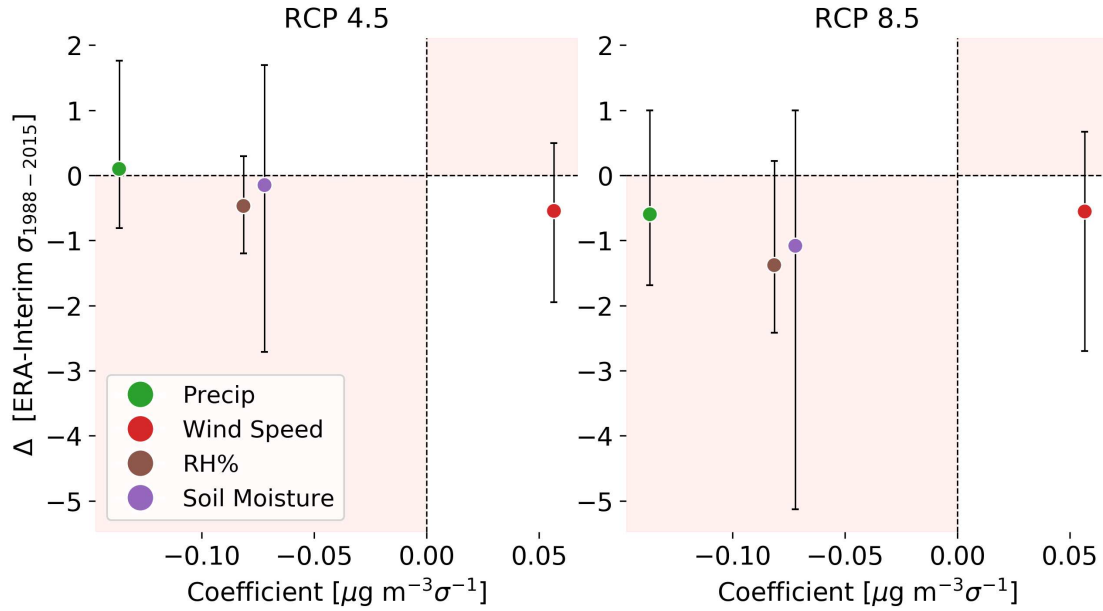
FIG. 4.4. Lasso regression coefficients vs. the the mean (dots) and total spread (min and max) of CMIP5 model changes from the historical to future period.

We use CMIP5 dust-season estimates of selected candidate variables to estimate future dust season fine dust concentrations using the relaxed-Lasso regression. The number of models available to show intermodel spread of regression estimated dust concentrations is reduced in this section due to the requirement that a single CMIP5 model must output all four selected candidate variables in order to estimate dust-season concentrations using the relaxed-Lasso model. The multimodel ensemble mean estimate of future dust concentrations, as well as the distribution of individual dust season concentrations estimated by individual CMIP5 models are shown in Figure 4.5. We estimate mean dust season concentrations to increase by 0.02 and 0.24 μg m^{-3} for RCP 4.5 and 8.5 respectively by 2080-2099. Some estimates of individual future dust season concentrations are much larger than values observed in the historical period, showing that although on average long term mean changes are small, dust seasons with significantly higher concentrations become more likely in the future. The multimodel ensemble-mean estimate in yearly dust concentrations show large variance in time compared to the single variable regression analysis in the previous section. For RCP 4.5 the range of values for the ensemble mean estimate is larger than the long-term mean increases. The boxplots in the bottom row show the distribution of single season estimates of dust concentrations for individual CMIP5 models. For the years 2080-2099 RCP 8.5, the majority of future dust season fine dust concentrations exceed the historical mean value (shown by the horizontal dashed line). Similar to the single variable regressions, changes

in future dust concentrations are on average larger for RCP 8.5 than 4.5. The range of dust season concentrations shown here is an underestimate as the regression fails to explain the observed historical variability and the historical data are unlikely to represent the true spread in southwest dust season concentrations.
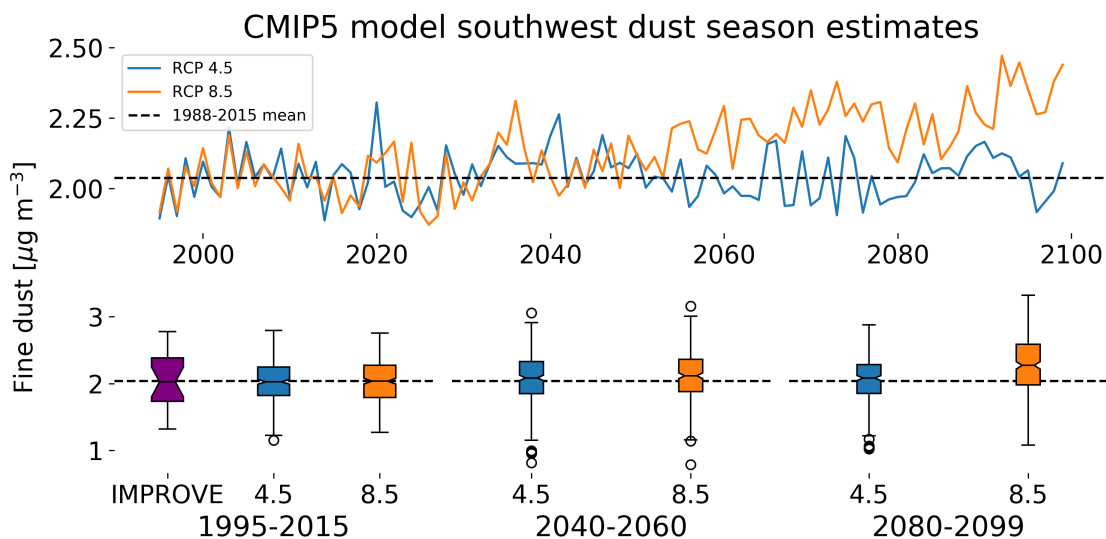

FIG. 4.5. Top row: Multi-model ensemble mean time series estimates of future dustiness estimated by Lasso regression where CMIP5 estimates of precipitation, wind speed, relative humidity, and soil moisture are the candidate variables used to drive the estimates. The blue and orange lines are the multi-model mean of RCP 4.5 and 8.5, respectively. Bottom row: Boxplots show the distribution of individual dust season concentrations estimated by each model and RCP scenario for the historical period (1988-2015), mid 21st century (2040-2060), and the end of the 20th century (2080-2099). The boxes extend from the lower to upper quartiles of the dust season concentration data, with the line in the middle showing the median value, whiskers extend to 1.5 times the interquartile range of the data. Values outside of this range are plotted as open circles. Note that the y-axis scale differ for the top and bottom rows. Scenario estimates use 16 and 15 models for RCP 4.5 and 8.5 respectively.

4.4 DISCUSSION

We show that estimates of future fine dust concentrations in the southwest are sensitive to the choice of environmental variables used to make the estimate, and that the differences in estimates based on different environmental variables can be larger than the differences observed between RCP 4.5 and 8.5. Of the variables that have been demonstrated to impact southwest dust concentrations, historical linear relationships show that increased temperature and VPD project the largest increases while changes in leaf area index suggest the greatest reduction. Changes estimated by a linear combination of candidate variables selected objectively do not use temperature, VPD, or LAI, and show a

more modest increase in dust concentrations. Our multilinear estimates of increased dust concentrations of 0.24 μg m^{-3} for RCP 8.5 by the end of the 21st century are similar to other findings. The difference of the mean from the historical period to the future period is statistically significant for RCP 8.5 (H0: μ_1-$\mu_2 = 0$, Ha: μ_1-$\mu_2 \neq 0$, p-value<0.001), the change for RCP 4.5 is not (p-value>0.001). Achakulwisut et al. (2018) estimate future southwest dust concentrations and estimate increases of 0.49 and 0.55 μg m^{-3} for March through May and June through August, respectively.

One key limitation of this work is the assumption that relationships between dust concentrations and environmental drivers will remain unchanged in the future. Another issue is the limited number of years available for the historical period, which limits our regression analysis approach by not fully capturing the phase space of dust concentrations and environmental condition values. Airborne dust can influence radiative budgets, and therefore alter temperature, precipitation, and other environmental variables that can influence dust concentrations. Our methods do not account for these types of feedbacks. Mean dust-season concentrations are impacted by episodic events. Though our work implies reduced dust emissions and concentrations with reductions in wind speed, other work has indicated that these more stagnant conditions may actually trap pollution and increase exposure if dust is present (Dawson et al. 2014).

In future work, additional historical variability may be explained by regression analysis if candidate variables were tailored geographically to the desert source regions, specifically to the Chihuahuan, Sonoran, and Mojave deserts. This geographic refinement would also help determine what specific source regions drive the dust concentrations observed over the southwest and help account for varying dust emission - environmental conditions relationships for different source regions. Another way to increase the fraction of historical variance explained may be to apply physically relevant transformations to candidate variables. For example, in this work we use LAI values as a proxy for the bareness of land. Evans et al. (2016) put forth a more explicit estimate of bareness using LAI, where bareness = exp(-1 x LAI). This method was used by Pu and Ginoux (2018) to estimate the influence of bareness on dust optical depth estimates made by CMIP5 models. Our work makes no attempt to estimate how fine dust concentrations may be influenced by long-range transport, quantifying this influence and how it may change in the future would make future work more robust. Dust emissions only occur when winds strong enough to loft dust occur over sufficiently dry dust sources (Fécan et al. 1998). This analysis is not set up to explicitly capture these events, which occur at submonthly timescales. However, relationships between monthly and daily conditions have been observed (*e.g.* Evan et al. (2016)) and many of

the monthly environmental conditions analyzed in this work can influence surface moisture, an important control on what wind speeds are required to loft dust. Future work could be improved by using daily data, though this would make robust intermodel comparisons more challenging.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND IDEAS FOR FUTURE RESEARCH

5.1 SUMMARY

This dissertation presents research towards a better understanding of the historical variability and possible futures for western U.S. PM emissions and concentrations from wildfires and dust. This work is separated into three papers. In the first paper (Chapter 2, published in *AGU Earth's Future*, Brey et al. (2018a)) I used the Fire Program Analysis Wildfire Occurrence Data (FPA FOD) to contrast the environmental conditions, meteorological drivers and air quality impacts of human- versus lightning-ignited wildfires in the southeast and western U.S. In the second paper (Chapter 3, submitted to *AGU Earth's Future*) I identified key environmental variables that best explain historical variability in summer wildfire burn area, quantified how these environmental predictors will change in the future, and estimated how simulated changes will alter future wildfire burn area. The third paper (Chapter 4, in preparation for *AGU Earth's Future*) examined how the spread in future southwest U.S. fine dust concentrations are sensitive to what environmental variable is used to predict dust concentrations.

5.1.1 Chapter 2 Overview

I found that the proportion of wildfires that are started by humans is higher in the southeast U.S. than in the west, though the seasonality of when these wildfires occur is also different (*i.e.*, the wildfires in the southeast are bimodal with more occurrences in the spring and fall months, whereas the majority of wildfires in the west occur during July/August). We show that there are larger contrasts in 1000-hour fuel-moisture between ecoregions than between ignition types, which implies that both ignition types are similarly constrained by fuel-moisture within a given ecoregion. Presently, both human- and lightning-ignited wildfire burn area are anti-correlated with total annual precipitation and will likely react similarly to future changes in precipitation (assuming a stationary relationship between precipitation and burn area). Between 1992 and 2015 humans were the dominant source of wildfire ignitions in the southeast U.S. However, the annual burn area of these wildfires is still linked to environmental conditions that allow fuels to ignite and wildfires to spread. Thus, climate change, not just human-ignited wildfires, will be an important driver of future wildfire activity and the resulting air quality impacts in the southeast U.S. The same is true for the west, where summertime burn area for both ignition types is greater in warmer, drier years. On average, wildfires in the southeast are smaller than in the west. However, these small wildfires significantly impact southeast air quality because 1) there is a large number

of southeast wildfires associated with National Weather Service air quality smoke forecasts, and 2) total $PM_{2.5}$ emissions from human-ignited wildfires in the southeast are similar to the total $PM_{2.5}$ emissions from human-ignited wildfires in the west.

5.1.2   *Chapter 3 overview*

Since the 1980s, the western U.S. has warmed and wildfire burn area has increased. Warm summers have been associated with a longer wildfire season and more burn area than cool summers. The Coupled Model Intercomparison Project phase 5 (CMIP5) models agree that western U.S. temperatures will continue to increase in the 21st century for representative concentration pathways (RCPs) 4.5 and 8.5. We find that averaged over seasonal and regional scales, other environmental variables demonstrated to be relevant to flammability, moisture abundances, and aridity such as precipitation, evaporation, relative humidity, root zone soil moisture, and wind speed can be used to explain observed variance in wildfire burn area as well or better than temperature. However, the magnitude and sign of the change of these variables in the 21st century are less certain than temperature. Temperature is the only variable that increases across all ecoregions, models, and RCPs. Our work demonstrates that when objectively selecting environmental variables to maximize predictive skill of linear regressions (minimize square error on unseen data) temperature is not always selected, but that this varies by western U.S. ecoregion. When temperature is not selected, the sign and magnitude of future changes in burn area become less certain. Future wildfire burn area estimates are lower whenever and wherever the importance of temperature is reduced in statistical models.

5.1.3   *Chapter 4 overview*

Chapter 4 examined how environmental conditions that drive fine dust emissions and concentrations in the southwestern U.S. change in the future and what these changes imply for future dust concentrations based on historical relationships. I examined environmental conditions identified by previous studies to influence dust emissions including, temperature, vapor pressure deficit, relative humidity, precipitation, soil moisture, wind speed, and leaf area index. My work quantifies fine dust concentrations in the U.S. southwest using IMPROVE measurements of fine iron as a proxy for dust, a method put forward by Hand et al. (2016). I averaged dust concentrations in the months March through July, the months with the highest concentrations, between 1995 and 2015. I showed that the largest contribution to the spread in future dust concentration estimates comes from the choice of variable used to explain observed variability in dust concentrations. The spread between variable estimates is larger than the spread between climate scenarios, or intermodel spread. The majority of models in the

fifth phase of the Coupled Model Intercomparison Project (CMIP5) that simulate leaf area index (LAI, a quantity anticorrelated with dust emissions and concentrations) show increasing leaf area index in the southwest U.S. throughout the 21st century. Based on our linear estimates of dust dependence on LAI, this LAI increase would result in reduced dust concentrations in the future. However, when I objectively selectrf environmental predictors of dust concentrations using Lasso regression, LAI was not selected in favor of other variables. When using a linear combination of the objectively selected environmental variables, I estimated that southwest U.S. future mean fine dust concentrations are expected to increase by 0.24 μg m^{-3} (12%) by the end of the 21st century for RCP 8.5. This estimated increase in fine dust concentration is driven by decreases in relative humidity, precipitation, soil moisture, and buffered by decreased wind speeds.

5.2 RECOMMENDATIONS FOR FUTURE RESEARCH

First I present incremental improvements to the work in this dissertation, then I suggest logical extensions and directions for future work.

5.2.1 Incremental Improvements to the work presented in this dissertation

This section documents steps that could be taken that may provide incremental improvements to some of the work and findings presented in this dissertation. These suggestions are not profound, they represent a running tally of work I would like to do if not for the practical confines of limited resources and time. These improvements primarily apply to Chapter 3, though some are general enough that they could be applied to any of the work presented in this dissertation. These incremental future steps are presented as a list.

- Implement new code with an improved model selection routine, which may help explain additional historical variance of year-to-year wildfire burn area or dust concentrations. This work uses a grid search technique to find the optimal value for α, the value that sets the amount of penalty for model complexity in Lasso regression. A similar method could be applied to data transformations and even the type of model used. For example, I could apply mathematical transformations to predictive features, perform a grid search across versions of transformations, and use the transformation that offers the lowest error on unseen data. The same could be done for the type of statistical model used. For example, we could see if Ridge regression, Lasso regression, or a non-linear model offers better performance, and analyze that model

instead. This would also provide an opportunity to make physically relevant combinations of variables, such as precipitation minus evaporation.

- Chapter 3 and 4 could be strengthened by using a longer time period of overlap between the European Center for Medium range Weather Forecasting (ECMWF) and CMIP5 output for bias correcting CMIP5 output. Presently, the overlap periods are 1983-2016 and 1995-2015 for Chapters 3 and 4 respectively.

- Test the sensitivity of the results presented in Chapter 3 and 4 to the grid spacing of the common grid selected for all gridded data (ECMWF and CMIP5).

- Wildfire burn area could be separated by land cover type and burn severity. Knowing more about how climate-burn area relationships vary by land cover type would render results in Chapter 3 more specific to PM concentrations, as burn area typically needs to be multiplied by a land-cover specific emission factor to calculate emissions.

- Examine additional CMIP5 ensemble members. All of the work presented in Chapters 3 and 4 relied on a single CMIP5 ensemble member (r1i1p1). Because many models were used, internal variability of the climate system was still accounted for, however, an even more robust measure of spread could be obtained with additional ensemble members.

- If linear regression is used to estimate future wildfire burn area, it could be improved by incorporating a component to the model that accounts for previously burned area. This physically relevant limitation to what can burn in the future is not accounted for by this work and would likely reduce the burn area estimated in Chapter 3. Other studies have incorporated a feedback like this into their work. Hurteau et al. (2019) show that accounting for fuel availability as well as flammability decreases projected increases in burn area. Leaf area index could also be used to estimate fuel availability.

- To help predict future wildfire burn area, this work could be strengthened by adding non-biophysical predictors that better account for human decisions and values regarding wildfires. A recent paper by Scheller et al. (2019) shows that changes in human-ignitions and wildfire suppression activities could significantly alter the amount of wildfire on the landscape in the future. It seems unlikely that the public will accept ever increasing burn area and increasingly devastating wildfires. It is possible a tipping point exists, where more funds are allocated for fuel treatments and suppression activities. A better understanding of historical tipping points may inform if this could happen again in the future.

- I have taken many practical steps towards ensuring the quality of the work presented in this dissertation. I routinely audit old pieces of code that analysis relies on (fresh eyes really help!) and track changes in code and figures using version control. However, bugs in code can be hard to find when the results look the way you expect and when only a single set of eyes are responsible for thousands of lines of code. Confidence in the work presented in this dissertation could be improved if the same care taken to catching typos in drafts of manuscripts was applied to the code that the writing depends on. Wilson et al. (2014) offer a set of best practices for scientific computing, some of which would be easy to integrate into a graduate students work flow. For example, code reviews where students get together to share code, explain functionality, and ask for suggestions on how code could be improved would help catch bugs. This could improve code the same way manuscripts are improved by review. Research shows that code reviews are the most cost effective way of finding bugs (Oram and Wilson 2010). Long term productivity gains could result as well, as an entire research group could coalesce towards the best implementations of common tasks, and knowledge would not be lost when a student or research scientist leaves a lab group. Using paired programming can result in increased productivity when solving particularly challenging problems (Oram and Wilson 2010). Paired programming is the practice where two people write code together, one person writes the code, the second provides real time feedback and focuses more on overall design and consistency issues. Tutorials on how to improve scientific computing workflows can be found at https://software-carpentry.org/.

Finally, a larger task, that could build upon the methods of Chapter 3, could be to analyze other seasons wildfire activity and drivers (not just summer). For example, in recent years the largest wildfires in California have occurred in Fall and Winter. It would be useful to know how those seasons are changing under climate change. Recent studies have shown that increases in wildfire burn area are likely to come from an expanding wildfire season.

5.2.2 Southeast U.S. wildfires

I believe southeast U.S. wildfires and how they may change in a changing climate is currently understudied. The analysis presented in Chapter 3, where I identify key wildfire predictors and examine how they change in the future, could be applied to the southeast U.S. Much of the southeast has fuel loading ($gC\,m^{-2}$) on par with or greater than western U.S. ecoregions (Urbanski et al. 2018). Brey et al. (2018a) showed that the total $PM_{2.5}$ emissions in the southeast are already large, despite the average

wildfire being much smaller than in the western U.S. Will climate change create future conditions conducive to so-called megafires occurring in the southeast? Even if the future climate of the southwest was conducive to the kinds of large wildfires that are observed in the west today, do non-biophysical controls, such as population density, road networks, use of prescribed fire, and terrain exclude the possibility of a future where wildfires look more like the west? Chiodi et al. (2018) show that the number of days with weather conditions that allow safe use of prescribed fire in the southeast will likely decrease in the future. Will reduced use of prescribed fire result in even greater fuel loads and larger wildfires?

5.2.3 *Understudied links between dust and wildfires in the western U.S.*

Aside from their relationships with PM pollution, the links between dust and wildfires may not be immediately obvious. However, these systems may already feedback on each other in such a way that working on both problems at the same time may be a natural fit for future research. For example, dust has been shown to play a more important role than temperature in the timing of snowmelt in the Rocky Mountains (Painter et al. 2018), which has been implicated in modulating western U.S. wildfire occurrence (*e.g.* Westerling (2016); Westerling et al. (2006)). Wildfires may impact dust emissions as well. The presence of vegetation can greatly reduce soil erosion by wind and can limit dust emissions (Marticorena and Bergametti 1995; Woodward et al. 2005; Pu and Ginoux 2017; Kim et al. 2017). Wildfires can reduce, alter, or eliminate vegetation from the landscape, exposing soil previously held in place by vegetation. In the future, as wildfires continue to occur and the climate warms and dries, it is possible that when vegetation is burned, the local climate is no longer suitable for the regrowth of that species, which could result in growth of a new species, or bare ground long term (Parks et al. 2018). An example of how wildfires and dust can be linked was recently featured in a story on National Public Radio (NPR). In Northern Nevada, ash and dust was observed to be blowing off of a field previously covered by sagebrush, but due to the 2018 Martin Fire, was now bare ground (Ahearn 2019).

5.2.4 *Seasonal Smoke forecasting*

There is a growing body of evidence that wildfire burn area will continue to grow in the future (*e.g.* Chapter 3 and references therein). These estimated increases in burn area will likely translate to some amount of air quality degradation. A high probability of increased future air quality impacts combined with a large body of evidence that shows smoke is harmful to human health lead me to believe that the most practical way to prepare for the future is to develop seasonal and sub-weekly smoke forecast products. It is time to learn how to live with smoke.

A smoke forecast with a lead time of several months, that estimates seasonal mean $PM_{2.5}$ concentrations would be beneficial in several key ways. Public health officials could adjust the allocation of health resources that may be required in greater abundances than normal (*e.g.* inhaler refills, on-call nurses). Public officials could encourage people to buy filters capable of reducing indoor $PM_{2.5}$ concentrations for their homes in advance of the smoke season, which may help reduce shortages that could occur during a smoke event if the forecast verifies. For example, in the event a significant smoke season is forecast, school districts could make sure that they have working indoor air filtration systems in place and prepare indoor activities as an alternative to outdoor recess. I developed a prototype of a seasonal smoke forecast for ATS 681 in the Spring semester of 2018. I used temperature, precipitation, and relative humidity from February, March, and April to estimate mean $PM_{2.5}$ concentrations for the months June through August in the western U.S. mountains (shown in Figure 5.1). The prototype was able to explain 26% of the observed variability in year-to-year mean $PM_{2.5}$ concentrations. The work did not separate smoke $PM_{2.5}$ from all $PM_{2.5}$, but instead relied on the fact that year-to-year western U.S. $PM_{2.5}$ variability is largely controlled by wildfire smoke (Liu et al. 2016b). The work presented in Chapter 3 shows that antecedent environmental conditions can explain year-to-year variability in wildfire burn area. The same machine learning methods used in that chapter could be adapted to make an operational seasonal smoke forecasting product.
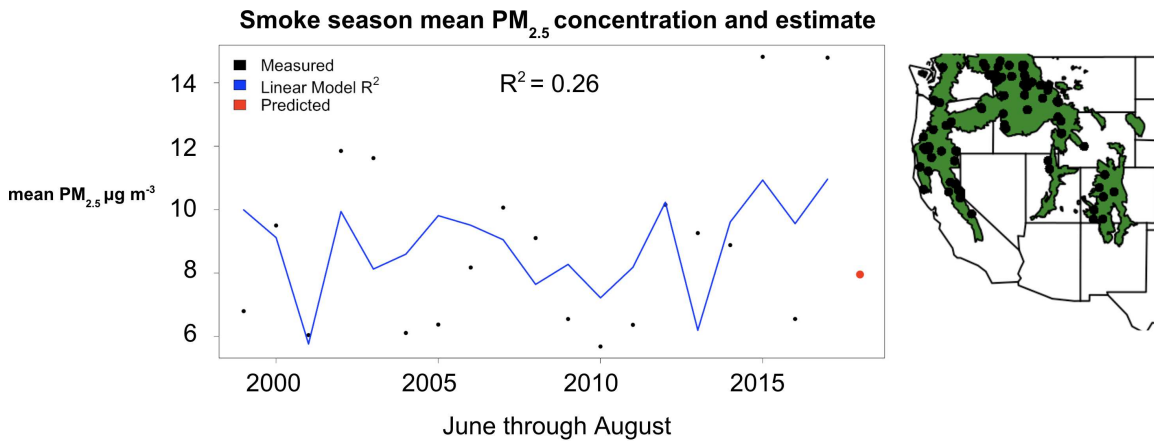

FIG. 5.1. Forested mountains (green area in map on right) mean June through August $PM_{2.5}$ concentrations (black dots) and linear estimate (blue line). The black dots on the map show the locations of monitors used to quantify $PM_{2.5}$. $PM_{2.5}$ data were downloaded from the U.S. EPA. Environmental conditions were quantified using mean Feb, March, April, Temperature, Precipitation, and Relative Humidity ECMWF ERA-Interim reanalysis.

There are several ways to try and improve the skill of the seasonal smoke forecast shown in Figure 5.1. 1) Wildfire activity could be forecast for specific source regions (*e.g.* Pacific northwest, Alberta). Once source specific estimates of wildfire activity are made, the smoke forecast could leverage the smoke transport climatologies developed by Brey et al. (2018b) to help estimate what downwind regions are most likely to be affected. Exceptionally vulnerable subregions could be identified by mapping low lying terrain vulnerable to smoke build up during inversions (at night), such as low lying valleys downslope of wildfire prone mountain ranges. This could be accomplished by identifying smoke risks on different hydrologic units (HUC) scales.

5.2.5 *Daily Smoke forecasts*

Estimates of smoke concentrations with lead times of 0 to 3 days could be useful as well. If disseminated to the public, people could change their planned activity in ways that would allow them to reduce their exposure to smoke. For example, soccer practice could be rescheduled or moved indoors, outdoor runs could be replaced by going to a gym, bike commutes could be replaced by taking the bus, *etc.*. Large audiences could be reached if T.V. meteorologist added a smoke forecast to their evening weather reports. A prototype of this kind of tool has already been developed by the United States Forest Service (USFS) AirFire team. They use the BlueSky smoke forecast model to estimate $PM_{2.5}$ concentrations up to 72 hours in the future (https://tools.airfire.org/websky/v1/). A team of researchers at Colorado State University lead by Dr. Marilee Long are currently combining the smoke forecast provided by the AirFire team with health impact assessment (HIA) research to build a website that forecast smoke concentrations as well as the expected health outcomes (*e.g.* expected number of emergency department visits due to smoke). Preliminary results for this work show that clearly communicating smoke forecasts, particularly the uncertainty and health outcomes remains a significant challenge. An annotated screenshot of this system, which I helped to build, is shown below in Figure 5.2.
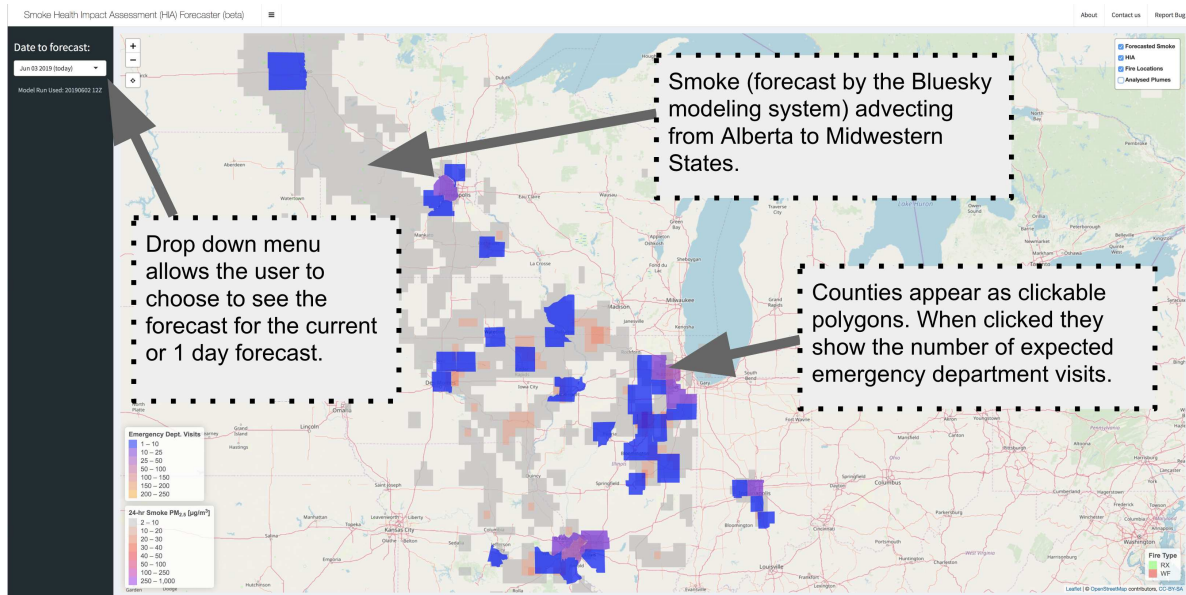
FIG. 5.2. Annotated screenshot of the operational prototype SmokeForecaster. The SmokeForecaster combines forecast smoke concentrations from BlueSky with HIA research to estimate the health burden of wildfire smoke. The screenshot shown above of http://rgan.atmos.colostate.edu was taken June 3rd 2019. The forecaster shows that smoke from wildfires in Alberta are impacting ground level $PM_{2.5}$ across several Midwestern States.

5.2.6   *Understanding how plants respond to climate change and how these changes may alter future wildfire activity*

NOTE: Many of the ideas and preliminary findings presented in this section would not be possible if not for the help of Dr. Abigail Swann. She revealed how plants changing physiology with increasing CO_2 could significantly impact how we estimate future wildfire occurrence. She pointed me to the idealized CMIP5 experiments required for the suggested future work presented here.

Wildfires primarily consume live and dead plants. Wind-blown dust is mostly emitted from surfaces that lack dense vegetation. The future of PM that result from both of these sources could be significantly altered by changes in the abundances, water use efficiency, or types of plants on the U.S. landscape in the future. Below is a short exploratory analysis focused on wildfires. I outline the ways in which plant physiology may impact wildfire occurrence and quantify how plants physiological responses to increasing CO_2 mixing ratios could impact future wildfire occurrence.

The most destructive wildfires consume both live and dead fuels (Jolly and Johnson 2018) and are influenced by fuel moisture content (Rothermel 1983; Simms and Law 1967). Living plants can reduce fire spread if they are not water stressed; as they typically have moisture content values an order of magnitude higher than dead fuels (Cohen et al. 1990). Even with equal fuel moisture content, live fuels

may be able to resist burning. Cohen et al. (1990) showed that attached (live) branches can absorb 49% more energy than detached branches, suggesting that plants can supply additional moisture to foliage when heated. In contrast, dead fuels passively respond to ambient atmospheric conditions. Changes in plant processes related to increased CO_2 mixing ratios could result in more fuels and more wildfire. Lasslop and Kloster (2015) estimate that since 1850, increased CO_2 mixing ratios has increased global wildfire emissions by 40%, due to increased fuel availability attributable to CO_2 fertilization.

Live fuels flammability and biomass are strongly influenced by plant physiology, thus plant physiology should be considered a critical component of the wildland fire system. Of potential importance to western U.S. wildfire occurrence, are the changes to plant physiology expected as CO_2 mixing ratios increase. These forecasted changes in plant physiology have been shown to impact environmental conditions correlated with wildfire occurrence (*e.g.* drought) (Swann et al. 2016). Plants absorb CO_2 through stomata in their leaves and lose water to the atmosphere through the same pathway (Swann et al. 2016). Some plants lose less water per unit of carbon gain when CO_2 mixing ratios increase because the gradient of CO_2 between the leaf and atmosphere is reduced (Cowan 1978). However, plant water use efficiency will only increase if the loss of water per unit of carbon gained is not offset by an increase in leaf area (Field et al. 1995). Assuming nearly constant leaf area, plant increased water use efficiency could result in reduced transpiration and increased soil moisture (Field et al. 1995). Combined, these changes could reduce plant water stress, even during droughts, and alter atmospheric conditions, like relative humidity. Observations show a decrease in transpiration due to increased water use efficiency (WUE) with increasing CO_2 mixing ratios (Keenan et al. 2013; Peñuelas et al. 2011; van der Sleen et al. 2014; Warren et al. 2011), though the number of plant species this applies to and limits are not fully understood (Battipaglia et al. 2013).

Estimates of the impact of increased CO_2 mixing ratios on plant physiology and water fluxes are integrated into the Earth System Models (ESMs) used in CMIP5 (Swann et al. 2016). In principal, some metrics that have been shown to be correlated with wildfire activity are not as strongly influenced by these processes as others. For example, potential evapotranspiration (PET) is the rate at which the atmosphere demands water from well-watered vegetation and is often estimated using the Penman-Monteith equation, which does not account for changes in plant physiology. Because PET does not account for how live fuels can regulate how much moisture is lost to the atmosphere differences in PET and actual transpiration can exist (Swann et al. 2016). These plant physiological processes can significantly alter our view of the impacts of climate change on events with severe consequences for humans.

For example, considering drought metrics that do not account for these plant physiological processes (*e.g.* Palmer Drought Severity Index, Standardized Precipitation Index) (Asadi Zarch et al. 2015; Cook et al. 2014), the prediction of future drought stress is much higher than studies that use metrics that better account for changes in plant physiological processes (*e.g.* Precipitation minus evaporation, soil moisture) (Koirala et al. 2014; Orlowsky and Seneviratne 2012; Roderick et al. 2015; Swann et al. 2016). Swann et al. (2016) categorizes environmental variables as plant-centric when they explicitly include the influence of atmospheric CO_2 on plant processes and evapotranspiration (*e.g.* evaporation, runoff, soil moisture, and relative humidity). Variables that do not account for these processes are considered to be atmospheric-centric.

The work in Chapter 3 identified the environmental variables that best explain historical year-to-year variability in wildfire burn area for specific ecoregions. The exploratory work presented next quantifies the impact of plant physiology on candidate environmental variables using output from two idealized CMIP5 experiments. In one experiment atmospheric CO_2 only influences atmospheric radiative transfer (CMIP5 experiment name esmFdbk1). In the second experiment the effect of CO_2 is isolated to plant physiology, such that CO_2 directly influences only photosynthetic processes (experiment name esmFixClim1). These experiments allow us to estimate the importance of plant physiology under changing CO_2 mixing ratios for variables output by CMIP5 models. The difference between variables for the mean of 20 years at high CO_2 mixing ratios (1,140 ppm) minus the mean of 20 years at low CO_2 (284 ppm) were estimated for each experiment, δ_{Rad} and δ_{Phys} respectively. I estimate the importance of plant physiology for a given variable with following equation:

$$Aphys_{abs} = \frac{|\delta_{Phys}|}{|\delta_{Phys} + \delta_{Rad}|}$$
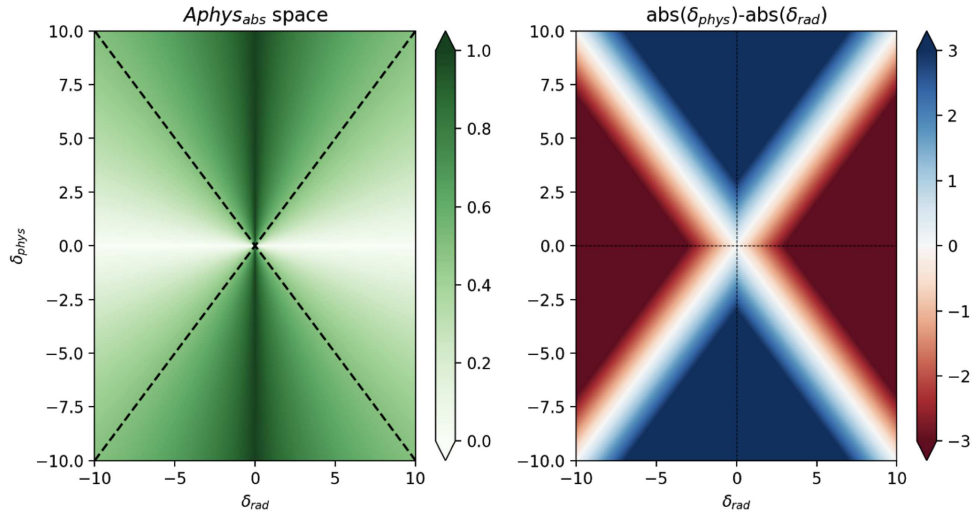
FIG. 5.3. Left: $Aphys_{abs}$ as a function of combinations of δ_{Phys} and δ_{Rad} values. Right: The difference between the absolute values of δ_{Phys} minus δ_{Rad}.

Aphys$_{abs}$ is estimated for each grid cell for individual variables amongst models that participated in the experiments. The minimum possible value is 0, the maximum is 1. When the value of Aphys$_{abs}$ is near 1 it indicates that changes to environmental variables for the experiment where only plants 'see' changing CO_2 are much larger than changes observed when only atmospheric radiative budgets 'see' changing CO_2. We interpret values near one as variables where changes are extremely sensitive to changing plant physiology with increasing CO_2. Values near 0 show changes are independent of plant physiology. The properties of Aphys$_{abs}$ as a function of δ_{Phys} and δ_{Rad} values are shown in Figure 5.3. We take the spatial mean of Aphys$_{abs}$ over the entire western U.S. for each CMIP5 model, variable, and season used in Chapter 3. The range of Aphys$_{abs}$ values across these subsets of data are shown in Figure 5.4.
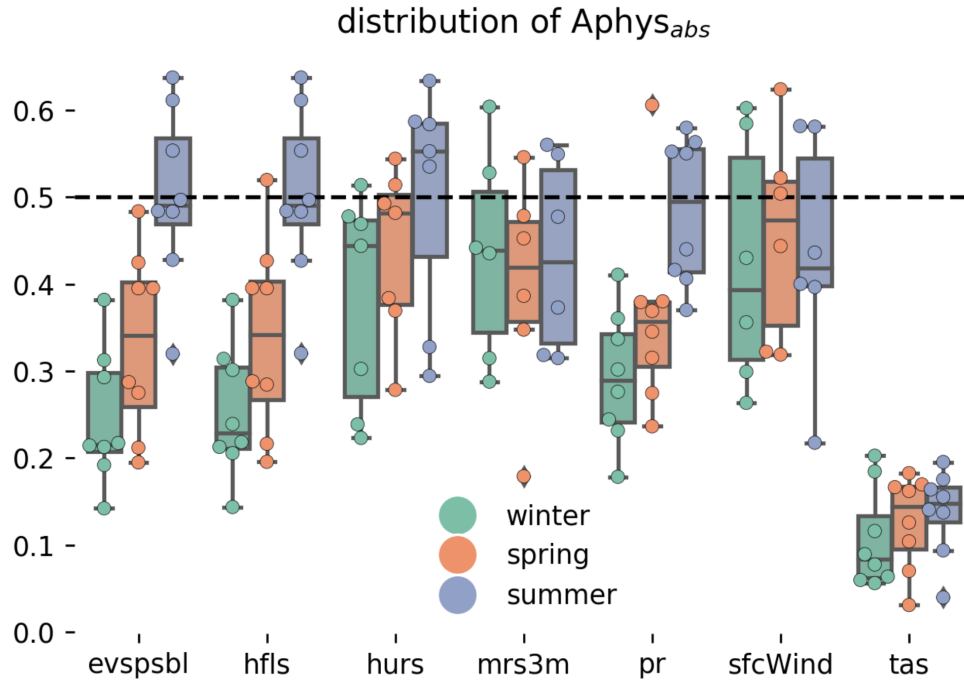
FIG. 5.4. The distribution of Aphys_{abs} over the western U.S. (3.1) for environmental variables (Evaporation (spsbl), latent heat flux (hfls), relative humidity (hurs), 3-meter soil moisture (mrs3m), precipitation (pr), surface wind (sfcWind), and temperature (tas)) averaged over spring (March, April, May), summer (June, July, August), and winter (November, December, January, February). Different variables are shown across the horizontal axis, different seasons are different colors, and individual model estimates are shown with dots over the boxplots. The black horizontal line at 0.5 indicates where the observed change in the experiment where only plants see changing CO_2 is as large as the experiment where only atmospheric radiative transfer responds to increased CO_2.

Figure 5.4 shows that changes in plant physiology due to increased CO_2 mixing ratios has a non-zero impact on all of the candidate variables presented in this analysis, indicating that plant physiology can alter the environmental conditions related to wildfire activity. Some candidate variables and some seasons are more strongly impacted than others. The dashed horizontal line in Figure 5.4 marks a significant boundary; above that line, according to the idealized CMIP5 experiments, plant physiology is more important for changes in variables than changes in the atmospheric radiative budget (though this is a very high bar to consider plant physiology important). All variables aside from temperature have at least one model in one season above that line. Each variable shows some amount of seasonality. Plant physiology impacts are largest in the summer and spring. Near surface relative humidity (hurs) is strongly influenced across seasons while near surface temperature (tas) is not. Figure 5.4 shows that

among the candidate variables chosen for this work, temperature is least affected by changes in plant physiology relative to changes in the atmospheric radiative budget for a given change in CO_2.
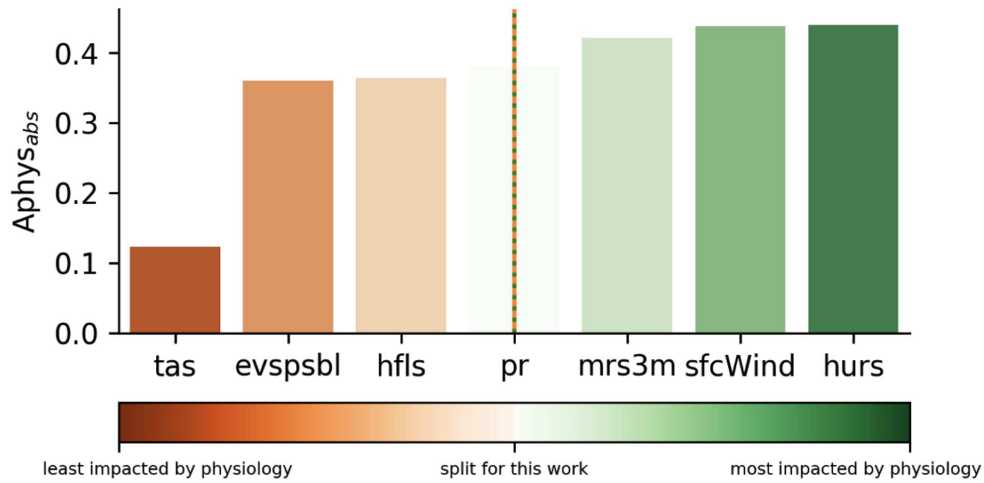

FIG. 5.5. The mean values of Aphys$_{abs}$ averaged across all seasons for each variable is shown in ascending order (least influenced by plant physiology to most influenced). The color scale matches the vertical axis and indicates the amount a variable is influenced by plant physiology. The vertical line in the center of the plot shows where the category for variables could be made for future work.

Figure 5.4 shows that there are no variables that are not influenced by plant physiology, and that spread among models and seasons can make ranking the degree of influence messy. Figure 5.5 shows the multi-season average of Aphys$_{abs}$ values for each variable. The difference in Aphys$_{abs}$ for variables in the middle of Figure 5.5 is small. Future work may gain insights on the impacts of plant-physiology on wildfire burn area by developing separate statistical models for the variables left and right of the vertical line shown in Figure 5.5.

Finally, I present changes in δ_{Phys} (changes in candidate variables attributed to plant physiology alone) multiplied with the sign of the regression coefficients for selected candidate variables from Chapter 3. When the sign of the regression coefficient matches the sign of the change from low CO_2 to high CO_2, the historical linear regressions developed in Chapter 3 suggest that plant physiology induced changes to that environmental variable would act to increase burn area. When the sign of the change and the regression are not the same, plant-physiology changes the environmental variable in a way that we would expect reduced burn area.
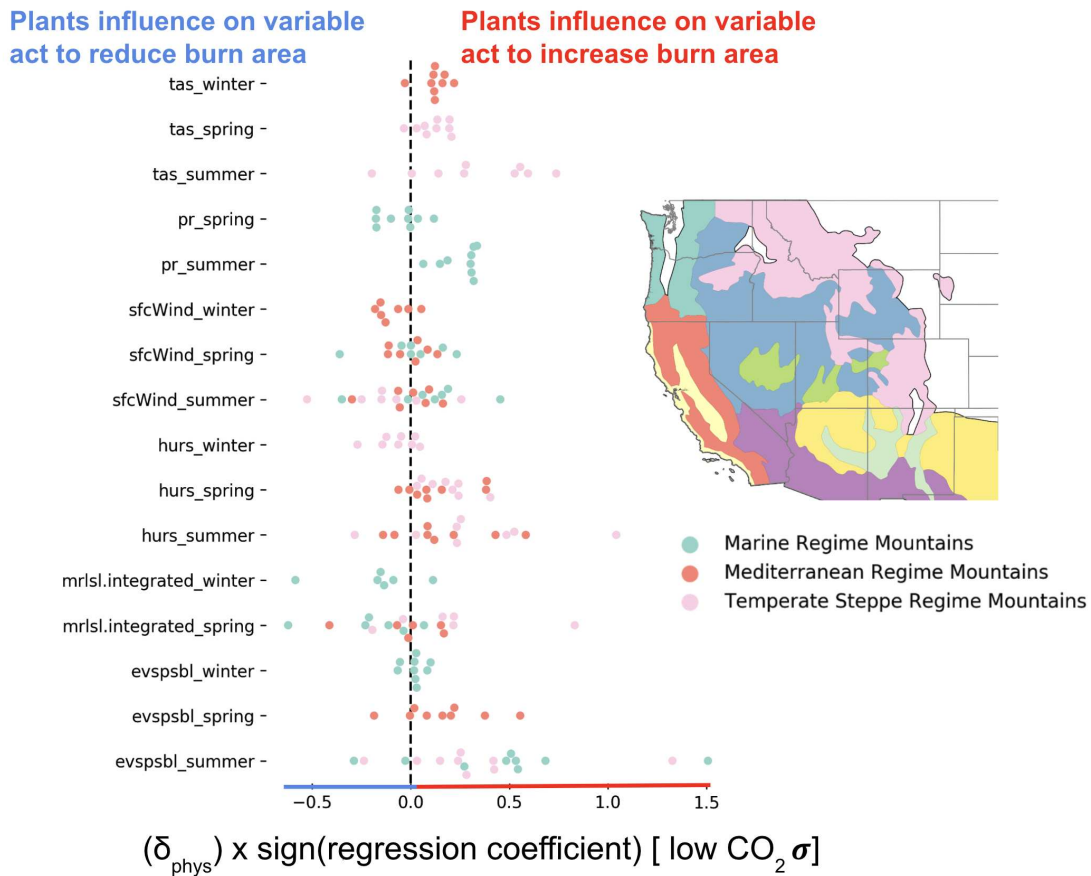
FIG. 5.6. The change in candidate variables estimated by individual CMIP5 models with output from the experiment where only plants saw changing CO_2 mixing ratios (esmFixClim1) multiplied by the sign of the regression coefficient from Chapter 3. Dots (models) right of zero represent models with changes from from low CO_2 to high CO_2 that have the same sign as the regression coefficient determined from the historical relationships between reanalysis data and burn area. Thus, dots with values greater than zero on the horizontal axis represent models that estimate changes that increase burn area in the context of the regressions developed in Chapter 3. The dots are color coded by ecoregion, shown in the inset map on the right. Only the mountainous ecoregions studied in Chapter 3 are presented here.

Figure 5.6 shows that CMIP5 models do not agree on the influence changing plant physiology will have on environmental variables shown to explain historical year-to-year variability in wildfire burn area. The change in variable values from low CO_2 to high CO_2 multiplied by regression coefficients from Chapter 3 shown in Figure 5.6 show that some CMIP5 models (dots in 5.6) estimate plant physiology will enhance future wildfire burn area (dots right of vertical line) while others may buffer increases (dots left of vertical line). The data presented in Figure 5.6 are preliminary and are only presented here as

motivation for future work in this area. The data presented in Figure 5.6 rely on δ_{Phys} values averaged over the entire western U.S. (spatial domain represented by all ecoregions shown in Figure 5.6). The results may change, if ecoregion specific δ_{Phys} values were multiplied by ecoregion specific regression coefficients. In future versions of Figure 5.6, it would be useful to better account for the magnitude of δ_{Phys}, not just the sign. Aphys$_{abs}$ describes how important plant physiology is relative to changes in the atmospheric radiative budget, but it does not show the magnitude of the changes in δ_{Phys} or δ_{rad}, which will be important to estimating impact of burn area.

References

Abatzoglou, J. T., 2013: Development of gridded surface meteorological data for ecological applications and modelling. *Int. J. Climatol.*, **33 (1)**, 121–131, doi: 10.1002/joc.3413.

Abatzoglou, J. T. and C. A. Kolden, 2013: Relationships between climate and macroscale area burned in the western united states. *Int. J. Wildland Fire*, **22 (7)**, 1003–1020, doi: 10.1071/WF13019.

Abatzoglou, J. T., C. A. Kolden, J. K. Balch, and B. A. Bradley, 2016: Controls on interannual variability in lightning-caused fire activity in the western US. *Environ. Res. Lett.*, **11 (4)**, 045 005, doi: 10.1088/1748-9326/11/4/045005.

Abatzoglou, J. T. and D. E. Rupp, 2017: Evaluating climate model simulations of drought for the northwestern united states: EVALUATING DROUGHT IN CLIMATE MODELS. *Int. J. Climatol.*, **37**, 910–920, doi: 10.1002/joc.5046.

Abatzoglou, J. T. and A. P. Williams, 2016: Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci. U. S. A.*, **113 (42)**, 11 770–11 775, doi: 10.1073/pnas.1607171113.

Achakulwisut, P., L. J. Mickley, and S. C. Anenberg, 2018: Drought-sensitivity of fine dust in the US southwest: Implications for air quality and public health under future climate change. *Environ. Res. Lett.*, **13 (5)**, 054 025, doi: 10.1088/1748-9326/aabf20.

Achakulwisut, P., L. Shen, and L. J. Mickley, 2017: What controls springtime fine dust variability in the western united states? investigating the 2002-2015 increase in fine dust in the U.S. southwest: Controlling factors of western U.S. dust. *J. Geophys. Res. D: Atmos.*, **122 (22)**, 12,449–12,467, doi: 10.1002/2017JD027208.

Achakulwisut, P., et al., 2019: Effects of increasing aridity on ambient dust and public health in the U.S. southwest under climate change. *GeoHealth*, doi: 10.1029/2019GH000187.

Ahearn, A., 2019: A sea of sagebrush disappears, making way for Fire-Prone cheatgrass. *NPR*.

Albini, F. A. and B. J. Stocks, 1986: Predicted and observed rates of spread of crown fires in immature jack pine.

Alexander, M. E. and M. G. Cruz, 2006: Evaluating a model for predicting active crown fire rate of spread using wildfire observations. *Can. J. For. Res.*, **36 (11)**, 3015–3028, doi: 10.1139/x06-174.

Allen, C. D., et al., 2010: A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *For. Ecol. Manage.*, **259**, 660–684, doi: 10.1016/j.foreco.2009.09.001.

Almukhtar, S., T. Griggs, K. Johnson, J. K. Patel, A. Singhvi, and D. Watkins, 2018: 'hell on earth': The first 12 hours of california's deadliest wildfire. *The New York Times*.

Anav, A., G. Murray-Tortarolo, P. Friedlingstein, S. Sitch, and Z. Zhu, 2013: Evaluation of land surface models in reproducing satellite derived leaf area index over the High-Latitude northern hemisphere. part II: Earth system models. *Remote Sensing*, **5 (8)**, 3637–3661, doi: 10.3390/rs5083637.

Anderson, H. E. and R. C. Rothermel, 1965: Influence of moisture and wind upon the characteristics of free-burning fires. *Symp. Combust.*, **10 (1)**, 1009–1019, doi: 10.1016/S0082-0784(65)80243-0.

Asadi Zarch, M. A., B. Sivakumar, and A. Sharma, 2015: Droughts in a warming climate: A global assessment of standardized precipitation index (SPI) and reconnaissance drought index (RDI). *J.*

Hydrol., **526**, 183–195, doi: 10.1016/j.jhydrol.2014.09.071.

Bachelet, D., T. Sheehan, K. Ferschweiler, and J. Abatzoglou, 2016: Simulating vegetation change, carbon cycling, and fire over the western united states using CMIP5 climate projections. *Natural Hazard Uncertainty Assessment*, John Wiley & Sons, Inc., 257–275, doi: 10.1002/9781119028116.ch17.

Bailey, R. G., 1983: Delineation of ecosystem regions. *Environ. Manage.*, **7 (4)**, 365–373, doi: 10.1007/BF01866919.

Baker, K. R., M. C. Woody, G. S. Tonnesen, W. Hutzell, H. O. T. Pye, M. R. Beaver, G. Pouliot, and T. Pierce, 2016: Contribution of regional-scale fire events to ozone and PM2.5 air quality estimated by photochemical modeling approaches. *Atmos. Environ.*, **140**, 539–554, doi: 10.1016/j.atmosenv.2016.06.032.

Balch, J. K., B. A. Bradley, J. T. Abatzoglou, R. C. Nagy, E. J. Fusco, and A. L. Mahood, 2017: Human-started wildfires expand the fire niche across the united states. *Proc. Natl. Acad. Sci. U. S. A.*, **114 (11)**, 2946–2951, doi: 10.1073/pnas.1617394114.

Barbero, R., J. T. Abatzoglou, E. A. Steel, and N. K. Larkin, 2014: Modeling very large-fire occurrences over the continental united states from weather and climate forcing. *Environ. Res. Lett.*, **9 (12)**, 124 009, doi: 10.1088/1748-9326/9/12/124009.

Battipaglia, G., M. Saurer, P. Cherubini, C. Calfapietra, H. R. McCarthy, R. J. Norby, and M. Francesca Cotrufo, 2013: Elevated co2 increases tree-level intrinsic water use efficiency: insights from carbon and oxygen isotope analyses in tree rings across three forest FACE sites. *New Phytol.*, **197 (2)**, 544–554, doi: 10.1111/nph.12044.

Berg, A., J. Sheffield, and P. C. D. Milly, 2017: Divergent surface and total soil moisture projections under global warming: Future soil moisture changes in coupled model intercomparison project phase 5. *Geophys. Res. Lett.*, **44 (1)**, 236–244, doi: 10.1002/2016GL071921.

Bond, T. C., et al., 2013: Bounding the role of black carbon in the climate system: A scientific assessment. *J. Geophys. Res. D: Atmos.*, **118 (11)**, 5380–5552, doi: 10.1002/jgrd.50171.

Bradshaw, L. S., J. E. Deeming, R. E. Burgan, and J. D. Cohen, 1984: The 1978 national Fire-Danger rating system: technical documentation. Tech. rep., Ogden, UT. doi: 10.2737/INT-GTR-169.

Brenner, J., 1991: Southern oscillation anomalies and their relationship to wildfire activity in florida. *Int. J. Wildland Fire*, **1 (1)**, 73–78, doi: 10.1071/wf9910073.

Brey, S. J., E. A. Barnes, J. R. Pierce, C. Wiedinmyer, and E. V. Fischer, 2018a: Environmental conditions, ignition type, and air quality impacts of wildfires in the southeastern and western U.S. *Earth's Future*, doi: 10.1029/2018EF000972.

Brey, S. J. and E. V. Fischer, 2016: Smoke in the city: How often and where does smoke impact summertime ozone in the united states? *Environ. Sci. Technol.*, **50 (3)**, 1288–1294, doi: 10.1021/acs.est.5b05218.

Brey, S. J., M. Ruminski, S. A. Atwood, and E. V. Fischer, 2018b: Connecting smoke plumes to sources using hazard mapping system (HMS) smoke and fire location data over north america. *Atmos. Chem. Phys.*, **18 (3)**, 1745–1761, doi: 10.5194/acp-18-1745-2018.

Brovkin, V., 2002: Climate-vegetation interaction. *J. Phys. IV*, **12 (10)**, 57–72, doi: 10.1051/jp4:20020452.

Bullard, J. E., S. P. Harrison, M. C. Baddock, N. Drake, T. E. Gill, G. McTainsh, and Y. Sun, 2011: Preferential dust sources: A geomorphological classification designed for use in global dust-cycle models. *J. Geophys. Res.*, **116 (F4)**, 17,069, doi: 10.1029/2011JF002061.

Cansler, C. A. and D. McKenzie, 2014: Climate, fire size, and biophysical setting control fire severity and spatial pattern in the northern cascade range, USA. *Ecol. Appl.*, **24 (5)**, 1037–1056.

Chen, L.-W. A., J. G. Watson, J. C. Chow, M. C. Green, D. Inouye, and K. Dick, 2012: Wintertime particulate pollution episodes in an urban valley of the western US: a case study. *Atmos. Chem. Phys.*, **12 (21)**, 10 051–10 064, doi: 10.5194/acp-12-10051-2012.

Chiodi, A. M., N. S. Larkin, and J. Morgan Varner, 2018: An analysis of southeastern US prescribed burn weather windows: seasonal variability and el niño associations. *Int. J. Wildland Fire*, **27 (3)**, 176–189, doi: 10.1071/WF17132.

Claiborn, C. S., D. Finn, T. V. Larson, and J. Q. Koenig, 2000: Windblown dust contributes to high PM25 concentrations. *J. Air Waste Manage. Assoc.*, **50 (8)**, 1440–1445, doi: 10.1080/10473289.2000.10464179.

Claverie, Martin, Eric Vermote, NOAA CDR Program (2014), 2014: NOAA climate data record (CDR) of leaf area index (LAI) and fraction of absorbed photosynthetically active radiation (FAPAR), version 4. *NOAA National Centers for Environmental Information*, doi: 10.7289/V5M043BX.

Cohen, W. B., P. N. Omi, and M. R. Kaufmann, 1990: Heating-Related water transport to intact lodgepole pine branches. *For. Sci.*, **36 (2)**, 246–254, doi: 10.1093/forestscience/36.2.246.

Collins, M., et al., 2013: Long-term climate change: Projections, commitments and irreversibility. *Climate Change 2013 - The Physical Science Basis*, T. F. Stocker, D. Qin, G.-K. Plattner, M. M. B. Tignor,

S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, Eds., Cambridge University Press, United Kingdom, Intergovernmental Panel on Climate Change, 1029–1136.

Collins, W. D., et al., 2006: Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the intergovernmental panel on climate change (IPCC) fourth assessment report (AR4). *J. Geophys. Res.*, **111**, D14 317.

Cook, B. I., J. E. Smerdon, R. Seager, and S. Coats, 2014: Global warming and 21st century drying. *Clim. Dyn.*, **43 (9)**, 2607–2627, doi: 10.1007/s00382-014-2075-y.

Correia, A. W., C. A. Pope, D. W. Dockery, Y. Wang, M. Ezzati, and F. Dominici, 2013: Effect of air pollution control on life expectancy in the united states an analysis of 545 US counties for the period from 2000 to 2007. *Epidemiology*, **24 (1)**, 23–31.

Cowan, I. R., 1978: Stomatal behaviour and environment. *Advances in Botanical Research*, R. D. Preston and H. W. Woolhouse, Eds., Academic Press, Vol. 4, 117–228, doi: 10.1016/S0065-2296(08)60370-5.

Crevoisier, C., E. Shevliakova, M. Gloor, C. Wirth, and S. Pacala, 2007: Drivers of fire in the boreal forests: Data constrained design of a prognostic model of burned area for use in dynamic global vegetation models. *J. Geophys. Res.*, **112 (D24)**, 512, doi: 10.1029/2006JD008372.

Daly, C., D. R. Conklin, and M. H. Unsworth, 2009: Local atmospheric decoupling in complex topography alters climate change impacts.

Dawson, J. P., B. J. Bloomer, D. A. Winner, and C. P. Weaver, 2014: Understanding the meteorological drivers of U.S. particulate matter concentrations in a changing climate. *Bull. Am. Meteorol. Soc.*, **95 (4)**, 521–532, doi: 10.1175/BAMS-D-12-00181.1.

Dee, D. P., et al., 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.*, **137 (656)**, 553–597, doi: 10.1002/qj.828.

Delfino, R. J., et al., 2009: The relationship of respiratory and cardiovascular hospital admissions to the southern california wildfires of 2003. *Occup. Environ. Med.*, **66 (3)**, 189–197, doi: 10.1136/oem.2008.041376.

DeMott, P. J., K. Sassen, M. R. Poellot, D. Baumgardner, D. C. Rogers, S. D. Brooks, A. J. Prenni, and S. M. Kreidenweis, 2003: African dust aerosols as atmospheric ice nuclei: AFRICAN DUST AEROSOLS AS ICE NUCLEI. *Geophys. Res. Lett.*, **30 (14)**, 251, doi: 10.1029/2003GL017410.

Dennison, P. E., S. C. Brewer, J. D. Arnold, and M. A. Moritz, 2014: Large wildfire trends in the western united states, 1984-2011: DENNISON ET. AL.; LARGE WILDFIRE TRENDS IN THE WESTERN US. *Geophys. Res. Lett.*, **41 (8)**, 2928–2933, doi: 10.1002/2014GL059576.

Eidenshink, J. C., B. Schwind, K. Brewer, Z.-L. Zhu, B. Quayle, and S. M. Howard, 2007: A project for monitoring trends in burn severity. *Fire Ecology*, **3 (1)**, 321, doi: 10.4996/fireecology.0301003.

Epa, U. S. and OAR, 2014: NAAQS table.

Evan, A. T., C. Flamant, M. Gaetani, and F. Guichard, 2016: The past, present and future of african dust. *Nature*, **531 (7595)**, 493–495, doi: 10.1038/nature17149.

Fécan, F., B. Marticorena, and G. Bergametti, 1998: Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas. *Ann. Geophys.*, **17 (1)**, 149–157, doi: 10.1007/s00585-999-0149-7.

Field, C. B., R. B. Jackson, and H. A. Mooney, 1995: Stomatal responses to increased CO2: implications from the plant to the global scale. *Plant Cell Environ.*, **18 (10)**, 1214–1225, doi: 10.1111/j.1365-3040.1995.tb00630.x.

Finco, M., B. Quayle, Y. Zhang, J. Lecker, K. A. Megown, and C. K. Brewer, 2012: Monitoring trends and burn severity (MTBS): monitoring wildfire activity for the past quarter century using landsat data. *In: Morin, Randall S.; Liknes, Greg C., comps. Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012; 2012 December 4-6; Baltimore, MD. Gen. Tech. Rep. NRS-P-105. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station.[CD-ROM]: 222-228.*, 222–228.

Fisk, W. J. and W. R. Chan, 2017: Health benefits and costs of filtration interventions that reduce indoor exposure to PM2.5 during wildfires. *Indoor Air*, **27 (1)**, 191–204, doi: 10.1111/ina.12285.

Flannigan, M. D., M. A. Krawchuk, W. J. de Groot, B. Mike Wotton, and L. M. Gowman, 2009: Implications of changing climate for global wildland fire. *www.publish.csiro.au/journals/ijwf International Journal of Wildland Fire*, **18**, 483–507, doi: 10.1071/WF08187.

Ford, B., M. Burke, W. Lassman, G. Pfister, and J. R. Pierce, 2017: Status update: is smoke on your mind? using social media to assess smoke exposure. *Atmos. Chem. Phys.*, **17 (12)**, 7541–7554, doi: 10.5194/acp-17-7541-2017.

Forkel, M., W. Dorigo, G. Lasslop, I. Teubner, E. Chuvieco, and K. Thonicke, 2017: A data-driven approach to identify controls on global fire activity from satellite and climate observations (SOFIA v1). *Geoscientific Model Development*, **10 (12)**, 4443–4476, doi: 10.5194/gmd-10-4443-2017.

Fosberg, M. A., J. E. Deeming, and Rocky Mountain Forest and Range Experiment Station (Fort Collins, Colo.), 1971: *Derivation of the one- and ten-hour timelag fuel moisture calculations for*

fire-danger rating. Rocky Mountain Forest and Range Experiment Station, Forest Service, U.S. Dept. of Agriculture.

Fusco, E. J., J. T. Abatzoglou, J. K. Balch, J. T. Finn, and B. A. Bradley, 2016: Quantifying the human influence on fire ignition across the western USA. *Ecol. Appl.*, **26 (8)**, 2388–2399, doi: 10.1002/eap.1395.

Gan, R. W., et al., 2017: Comparison of wildfire smoke estimation methods and associations with cardiopulmonary-related hospital admissions. *Geohealth*, **1 (3)**, 122–136, doi: 10.1002/2017GH000073.

Gannet Hallar, A., et al., 2017: Impacts of increasing aridity and wildfires on aerosol loading in the intermountain western US. *Environ. Res. Lett.*, **12 (1)**, 014 006, doi: 10.1088/1748-9326/aa510a.

Giglio, L., J. T. Randerson, and G. R. van der Werf, 2013: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *J. Geophys. Res. Biogeosci.*, **118 (1)**, 317–328, doi: 10.1002/jgrg.20042.

Ginoux, P., M. Chin, I. Tegen, J. M. Prospero, B. Holben, O. Dubovik, and S.-J. Lin, 2001: Sources and distributions of dust aerosols simulated with the GOCART model. *J. Geophys. Res.*, **106 (D17)**, 20 255–20 273, doi: 10.1029/2000JD000053.

Grineski, S. E., J. G. Staniswalis, P. Bulathsinhala, Y. Peng, and T. E. Gill, 2011: Hospital admissions for asthma and acute bronchitis in el paso, texas: do age, sex, and insurance status modify the effects of dust and low wind events? *Environ. Res.*, **111 (8)**, 1148–1155, doi: 10.1016/j.envres.2011.06.007.

Hand, J. L., W. H. White, K. A. Gebhart, N. P. Hyslop, T. E. Gill, and B. A. Schichtel, 2016: Earlier onset of the spring fine dust season in the southwestern united states: Early onset of the spring SW dust season. *Geophys. Res. Lett.*, **43 (8)**, 4001–4009, doi: 10.1002/2016GL068519.

Harris, R. M. B., T. A. Remenyi, G. J. Williamson, N. L. Bindoff, and D. M. J. S. Bowman, 2016: Climate-vegetation-fire interactions and feedbacks: trivial detail or major barrier to projecting the future of the earth system? *WIREs Clim Change*, **7 (6)**, 910–931, doi: 10.1002/wcc.428.

Higuera, P. E., J. T. Abatzoglou, J. S. Littell, and P. Morgan, 2015: The changing strength and nature of Fire-Climate relationships in the northern rocky mountains, U.S.A., 1902-2008. *PLoS One*, **10 (6)**, e0127 563, doi: 10.1371/journal.pone.0127563.

Holden, Z. A., et al., 2018: Decreasing fire season precipitation increased recent western US forest wildfire activity. *Proc. Natl. Acad. Sci. U. S. A.*, **115 (36)**, E8349–E8357, doi: 10.1073/pnas.1802316115.

Hurteau, M. D., S. Liang, A. L. Westerling, and C. Wiedinmyer, 2019: Vegetation-fire feedback reduces projected area burned under climate change. *Sci. Rep.*, **9 (1)**, 2838, doi: 10.1038/s41598-019-39284-1.

Hurteau, M. D., A. L. Westerling, C. Wiedinmyer, and B. P. Bryant, 2014: Projected effects of climate and development on california wildfire emissions through 2100. *Environ. Sci. Technol.*, **48 (4)**, 2298–2304, doi: 10.1021/es4050133.

Hyslop, N. P., K. Trzepla, and W. H. White, 2015: Assessing the suitability of historical PM(2.5) element measurements for trend analysis. *Environ. Sci. Technol.*, **49 (15)**, 9247–9255, doi: 10.1021/acs.est.5b01572.

Jaffe, D., W. Hafner, D. Chand, A. L. Westerling, and D. Spracklen, 2008: Interannual variations in PM2.5 due to wildfires in the western united states. *Environ. Sci. Technol.*, **42 (8)**, 2812–2818.

Jaffe, D., N. Wigder, N. Downey, G. Pfister, A. Boynard, and S. B. Reid, 2013: Impact of wildfires on ozone exceptional events in the western US. *Environ. Sci. Technol.*, **47 (19)**, 11 065–11 072.

Jensen, D., J. T. Reager, B. Zajic, N. Rousseau, M. Rodell, and E. Hinkley, 2018: The sensitivity of US wildfire occurrence to pre-season soil moisture conditions across ecosystems. *Environ. Res. Lett.*, **13 (1)**, doi: 10.1088/1748-9326/aa9853.

Jolly, W. M. and D. M. Johnson, 2018: Pyro-Ecophysiology: Shifting the paradigm of live wildland fuel research. *Fire*, **1 (1)**, 8, doi: 10.3390/fire1010008.

Kashian, D. M., W. H. Romme, D. B. Tinker, M. G. Turner, and M. G. Ryan, 2006: Carbon storage on landscapes with stand-replacing fires. *Bioscience*, **56 (7)**, 598–606, doi: 10.1641/0006-3568(2006)56[598:CSOLWS]2.0.CO;2.

Keenan, T. F., D. Y. Hollinger, G. Bohrer, D. Dragoni, J. W. Munger, H. P. Schmid, and A. D. Richardson, 2013: Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise. *Nature*, **499 (7458)**, 324–327, doi: 10.1038/nature12291.

Keywood, M., et al., 2013: Fire in the air: Biomass burning impacts in a changing climate. *Crit. Rev. Environ. Sci. Technol.*, **43 (1)**, 40–83, doi: 10.1080/10643389.2011.604248.

Kim, D., M. Chin, L. A. Remer, T. Diehl, H. Bian, H. Yu, M. E. Brown, and W. R. Stockwell, 2017: Role of surface wind and vegetation cover in multi-decadal variations of dust emission in the sahara and sahel. *Atmos. Environ.*, **148**, 282–296, doi: 10.1016/j.atmosenv.2016.10.051.

Kloster, S. and G. Lasslop, 2017: Historical and future fire occurrence (1850 to 2100) simulated in CMIP5 earth system models. *Glob. Planet. Change*, **150**, 58–69, doi: 10.1016/j.gloplacha.2016.12.017.

Koirala, S., Y. Hirabayashi, R. Mahendran, and S. Kanae, 2014: Global assessment of agreement among streamflow projections using CMIP5 model outputs. *Environ. Res. Lett.*, **9 (6)**, 064 017, doi: 10.1088/1748-9326/9/6/064017.

Krueger, B. J., V. H. Grassian, J. P. Cowin, and A. Laskin, 2004: Heterogeneous chemistry of individual mineral dust particles from different dust source regions: the importance of particle mineralogy. *Atmos. Environ.*, **38 (36)**, 6253–6261, doi: 10.1016/j.atmosenv.2004.07.010.

Lasslop, G. and S. Kloster, 2015: Impact of fuel variability on wildfire emission estimates. *Atmos. Environ.*, **121**, 93–102, doi: 10.1016/j.atmosenv.2015.05.040.

Lassman, W., B. Ford, R. W. Gan, G. Pfister, S. Magzamen, E. V. Fischer, and J. R. Pierce, 2017: Spatial and temporal estimates of population exposure to wildfire smoke during the washington state 2012 wildfire season using blended model, satellite, and in situ data. *GeoHealth*, **1 (3)**, 2017GH000 049, doi: 10.1002/2017GH000049.

Lee, J. A., M. C. Baddock, M. J. Mbuh, and T. E. Gill, 2012: Geomorphic and land cover characteristics of aeolian dust sources in west texas and eastern new mexico, USA. *Aeolian Research*, **3 (4)**, 459–466, doi: 10.1016/j.aeolia.2011.08.001.

Lee, J. A., T. E. Gill, K. R. Mulligan, M. Dominguez Acosta, and A. E. Perez, 2009: Land use/land cover and point sources of the 15 december 2003 dust storm in southwestern north america. *Geomorphology*, **105 (1)**, 18–27, doi: 10.1016/j.geomorph.2007.12.016.

Li, R., C. Wiedinmyer, K. R. Baker, and M. P. Hannigan, 2013: Characterization of coarse particulate matter in the western united states: a comparison between observation and modeling. 1311–1327 pp., doi: 10.5194/acp-13-1311-2013.

Lipner, E. M., K. O'Dell, S. J. Brey, B. Ford, J. R. Pierce, E. V. Fischer, and J. L. Crooks, 2019: The associations between clinical respiratory outcomes and ambient wildfire smoke exposure among pediatric asthma patients at national jewish health, 2012-2015. *GeoHealth*, doi: 10.1029/2018GH000142.

Littell, J. S., D. McKenzie, D. L. Peterson, and A. L. Westerling, 2009: Climate and wildfire area burned in western U.S. ecoprovinces, 1916–2003. *Ecol. Appl.*, **19 (4)**, 1003–1021, doi: 10.1890/07-1183.1.

Littell, J. S., D. L. Peterson, K. L. Riley, Y. Liu, and C. H. Luce, 2016: A review of the relationships between drought and forest fire in the united states. *Glob. Chang. Biol.*, **22 (7)**, 2353–2369, doi: 10.1111/gcb.13275.

Liu, J. C., L. J. Mickley, M. P. Sulprizio, X. Yue, R. D. Peng, F. Dominici, and M. L. Bell, 2016a: Future respiratory hospital admissions from wildfire smoke under climate change in the western US. *Environ. Res. Lett.*, **11 (12)**, 124 018.

Liu, J. C., et al., 2016b: Particulate air pollution from wildfires in the western US under climate change. *Clim. Change*, 1–12, doi: 10.1007/s10584-016-1762-6.

Liu, Y., S. L. Goodrick, and J. A. Stanturf, 2013: Future U.S. wildfire potential trends projected using a dynamically downscaled climate change scenario. *For. Ecol. Manage.*, **294**, 120–135, doi: 10.1016/j.foreco.2012.06.049.

Loukina, A., K. Zechner, L. Chen, and M. Heilman, 2015: Feature selection for automated speech scoring. *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications*, 12–19, doi: 10.3115/v1/W15-0602.

Lu, X., L. Zhang, X. Yue, J. Zhang, D. A. Jaffe, A. Stohl, Y. Zhao, and J. Shao, 2016: Wildfire influences on the variability and trend of summer surface ozone in the mountainous western united states.

Atmos. Chem. Phys. Disc., 1–40, doi: 10.5194/acp-2016-646.

Mahowald, N., F. Lo, Y. Zheng, L. Harrison, C. Funk, D. Lombardozzi, and C. Goodale, 2016: Projections of leaf area index in earth system models. *Earth System Dynamics*, **7 (1)**, 211–229, doi: 10.5194/esd-7-211-2016.

Mahowald, N. M., et al., 2010: Observed 20th century desert dust variability: impact on climate and biogeochemistry. *Atmos. Chem. Phys.*, **10 (22)**, 10 875–10 893, doi: 10.5194/acp-10-10875-2010.

Malek, E., T. Davis, R. S. Martin, and P. J. Silva, 2006: Meteorological and environmental aspects of one of the worst national air pollution episodes (january, 2004) in logan, cache valley, utah, USA. *Atmos. Res.*, **79 (2)**, 108–122, doi: 10.1016/j.atmosres.2005.05.003.

Malm, W. C., J. F. Sisler, D. Huffman, R. A. Eldred, and T. A. Cahill, 1994: Spatial and seasonal trends in particle concentration and optical extinction in the united states. *J. Geophys. Res.*, **99 (D1)**, 1347, doi: 10.1029/93JD02916.

Marlon, J. R., et al., 2012: Long-term perspective on wildfires in the western USA. *Proc. Natl. Acad. Sci. U. S. A.*, **109 (9)**, E535–E543, doi: 10.1073/pnas.1112839109.

Marticorena, B. and G. Bergametti, 1995: Modeling the atmospheric dust cycle: 1. design of a soil-derived dust emission scheme. *J. Geophys. Res.*, **100 (D8)**, 16 415, doi: 10.1029/95JD00690.

McBride, A., 2018: Camp fire: Death toll rises to 86 after hospitalized man dies from burn injuries. San Francisco Chronicle, accessed: 2019-5-29, https://www.sfchronicle.com/california-wildfires/article/Camp-Fire-Death-toll-rises-to-86-after-13458956.php.

McClure, C. D. and D. A. Jaffe, 2018a: Investigation of high ozone events due to wildfire smoke in an urban area. *Atmos. Environ.*, **194**, 146–157, doi: 10.1016/j.atmosenv.2018.09.021.

McClure, C. D. and D. A. Jaffe, 2018b: US particulate matter air quality improves except in wildfire-prone areas. *Proc. Natl. Acad. Sci. U. S. A.*, **115 (31)**, 7901–7906, doi: 10.1073/pnas.1804353115.

McClure, C. D. and D. A. Jaffe, 2018c: US particulate matter air quality improves except in wildfire-prone areas. *Proc. Natl. Acad. Sci. U. S. A.*, **115 (31)**, 7901–7906, doi: 10.1073/pnas.1804353115.

Mcgill, R., J. W. Tukey, and W. A. Larsen, 1978: Variations of box plots. *Am. Stat.*, **32 (1)**, 12–16, doi: 10.1080/00031305.1978.10479236.

McKendry, I. G., A. Christen, S.-C. Lee, M. Ferrara, K. B. Strawbridge, N. O'Neill, and A. Black, 2019: Impacts of an intense wildfire smoke episode on surface radiation, energy and carbon fluxes in southwestern british columbia, canada. *Atmos. Chem. Phys.*, **19 (2)**, 835–846, doi: 10.5194/acp-19-835-2019.

McKENZIE, D., Z. Gedalof, D. L. Peterson, and P. Mote, 2004: Climatic change, wildfire, and conservation. *Conserv. Biol.*, **18 (4)**, 890–902, doi: 10.1111/j.1523-1739.2004.00492.x.

Meinshausen, N., 2007: Relaxed lasso. *Comput. Stat. Data Anal.*, **52 (1)**, 374–393, doi: 10.1016/j.csda.2006.12.019.

Melvin, M., 2015: National prescribed fire use survey report. Tech. rep., Coalition of Prescribed Fire Councils.

Miller, R. L. and I. Tegen, 1998: Climate response to soil dust aerosols. *J. Clim.*, **11 (12)**, 3247–3267, doi: 10.1175/1520-0442(1998)011<3247:CRTSDA>2.0.CO;2.

Mitchell, R. J., Y. Liu, J. J. O'Brien, K. J. Elliott, G. Starr, C. F. Miniat, and J. K. Hiers, 2014: Future climate and fire interactions in the southeastern region of the united states. *For. Ecol. Manage.*, **327**, 316–326, doi: 10.1016/j.foreco.2013.12.003.

Mora, C., et al., 2013: The projected timing of climate departure from recent variability. *Nature*, **502 (7470)**, 183–187, doi: 10.1038/nature12540.

Moritz, M. A., M. A. Parisien, E. Batllori, M. A. Krawchuk, J. Van Dorn, D. J. Ganz, and K. Hayhoe, 2012: Climate change and disruptions to global fire activity. *Ecosphere*, **3 (6)**.

Morton, D. C., G. J. Collatz, D. Wang, J. T. Randerson, L. Giglio, and Y. Chen, 2013: Satellite-based assessment of climate controls on US burned area. *Biogeosciences*, **10 (1)**, 247–260, doi: 10.5194/bg-10-247-2013.

Murph, Z. and C. Mooney, 2019: Montana's forests have swung from pulling carbon dioxide out of the air to putting it back again. *The Washington Post*.

Nagy, R. C., E. Fusco, B. Bradley, J. T. Abatzoglou, and J. Balch, 2018: Human-Related ignitions increase the number of large wildfires across U.S. ecoregions. *Fire*, **1 (1)**, 4, doi: 10.3390/fire1010004.

NCDC, ????: Did you know? Accessed: 2018-08-08, `https://www.ncdc.noaa.gov/monitoring-references/dyk/deadfuelmoisture`.

O'Dell, K., B. Ford, E. V. Fischer, and J. R. Pierce, 2019: Contribution of Wildland-Fire smoke to US PM2.5 and its influence on recent trends. *Environ. Sci. Technol.*, **53 (4)**, 1797–1804, doi: 10.1021/acs.est.8b05430.

Omernik, J. M., 1987: Ecoregions of the conterminous united states. *Ann. Assoc. Am. Geogr.*, **77 (1)**, 118–125, doi: 10.1111/j.1467-8306.1987.tb00149.x.

Omernik, J. M., 1995: Ecoregions: a spatial framework for environmental management. *Biological assessment and criteria: tools for water resource planning and decision making*, 49–62.

Oram, A. and G. Wilson, 2010: *Making Software: What Really Works, and Why We Believe It.* "O'Reilly Media, Inc.".

Orlowsky, B. and S. I. Seneviratne, 2012: Elusive drought: uncertainty in observed trends and short- and long-term CMIP5 projections. doi: 10.5194/hessd-9-13773-2012.

Painter, T. H., S. M. Skiles, J. S. Deems, W. T. Brandt, and J. Dozier, 2018: Variation in rising limb of colorado river snowmelt runoff hydrograph controlled by dust radiative forcing in snow. *Geophys. Res. Lett.*, **45 (2)**, 797–808, doi: 10.1002/2017GL075826.

Parisien, M.-A., C. Miller, S. A. Parks, E. R. DeLancey, F.-N. Robinne, and M. D. Flannigan, 2016: The spatially varying influence of humans on fire probability in north america. *Environ. Res. Lett.*, **11 (7)**, 075 005, doi: 10.1088/1748-9326/11/7/075005.

Park Williams, A., B. I. Cook, J. E. Smerdon, D. A. Bishop, R. Seager, and J. S. Mankin, 2017: The 2016 southeastern U.S. drought: An extreme departure from centennial wetting and cooling. *J. Geophys. Res. D: Atmos.*, **122 (20)**, 2017JD027 523, doi: 10.1002/2017JD027523.

Park Williams, A., et al., 2012: Temperature as a potent driver of regional forest drought stress and tree mortality. *Nat. Clim. Chang.*, **3**, 292, doi: 10.1038/nclimate1693.

Park Williams, A., et al., 2015: Correlations between components of the water balance and burned area reveal new insights for predicting forest fire area in the southwest united states. *Int. J. Wildland Fire*, **24 (1)**, 14–26, doi: 10.1071/WF14023.

Parks, S. A., L. M. Holsinger, C. Miller, and M.-A. Parisien, 2018: Analog-based fire regime and vegetation shifts in mountainous regions of the western US. *Ecography*, **41 (6)**, 910–921, doi: 10.1111/ecog.03378.

Pechony, O. and D. T. Shindell, 2010: Driving forces of global wildfires over the past millennium and the forthcoming century. *Proc. Natl. Acad. Sci. U. S. A.*, **107 (45)**, 19 167–19 170, doi: 10.1073/pnas.1003669107.

Pedregosa, F., et al., 2011: Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, **12 (Oct)**, 2825–2830.

Pendergrass, A. G., R. Knutti, F. Lehner, C. Deser, and B. M. Sanderson, 2017: Precipitation variability increases in a warmer climate. *Sci. Rep.*, **7 (1)**, 17 966, doi: 10.1038/s41598-017-17966-y.

Peñuelas, J., J. G. Canadell, and R. Ogaya, 2011: Increased water-use efficiency during the 20th century did not translate into enhanced tree growth: Tree growth in the 20th century. *Glob. Ecol. Biogeogr.*, **20 (4)**, 597–608, doi: 10.1111/j.1466-8238.2010.00608.x.

Peterman, W., D. Bachelet, K. Ferschweiler, and T. Sheehan, 2014: Soil depth affects simulated carbon and water in the MC2 dynamic global vegetation model. *Ecol. Modell.*, **294**, 84–93, doi: 10.1016/j.ecolmodel.2014.09.025.

Pope, C. A., 3rd, R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, and G. D. Thurston, 2002: Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA*, **287 (9)**, 1132–1141.

Prospero, J. M., 2002: Environmental characterization of global sources of atmospheric soil dust identified with the NIMBUS 7 total ozone mapping spectrometer (TOMS) absorbing aerosol product. doi: 10.1029/2000rg000095.

Prüss-Ustün, A., C. Vickers, P. Haefliger, and R. Bertollini, 2011: Knowns and unknowns on burden of disease due to chemicals: a systematic review. *Environ. Health*, **10**, 9, doi: 10.1186/1476-069X-10-9.

Pu, B. and P. Ginoux, 2017: Projection of american dustiness in the late 21st century due to climate change. *Sci. Rep.*, **7 (1)**, 5553, doi: 10.1038/s41598-017-05431-9.

Pu, B. and P. Ginoux, 2018: How reliable are CMIP5 models in simulating dust optical depth? *Atmos. Chem. Phys.*, **18 (16)**, 12 491–12 510, doi: 10.5194/acp-18-12491-2018.

Randerson, J. T., Y. Chen, G. R. van der Werf, B. M. Rogers, and D. C. Morton, 2012: Global burned area and biomass burning emissions from small fires. *J. Geophys. Res.*, **117 (G4)**, G04 012, doi: 10.1029/2012JG002128.

Rappold, A. G., et al., 2011: Peat bog wildfire smoke exposure in rural north carolina is associated with cardiopulmonary emergency department visits assessed through syndromic surveillance. *Environ. Health Perspect.*, **119 (10)**, 1415–1420, doi: 10.1289/ehp.1003206.

Reyes-Velarde, A., 2019: California's camp fire was the costliest global disaster last year, insurance report shows. *Los Angeles Times.*

Riley, K. L., J. T. Abatzoglou, I. C. Grenfell, A. E. Klene, and F. A. Heinsch, 2013: The relationship of large fire occurrence with drought and fire danger indices in the western USA, 1984–2008: the role of temporal scale. *Int. J. Wildland Fire,* **22 (7)**, 894–909, doi: 10.1071/WF12149.

Roderick, M. L., P. Greve, and G. D. Farquhar, 2015: On the assessment of aridity with changes in atmospheric CO2. *Water Resour. Res.,* **51 (7)**, 5450–5463.

Rolph, G. D., et al., 2009: Description and verification of the NOAA smoke forecasting system: The 2007 fire season. *Weather Forecast.,* **24 (2)**, 361–378, doi: 10.1175/2008WAF2222165.1.

Rosenfeld, D. and R. Nirel, 1996: Seeding effectiveness: Interaction of desert dust and the southern margins of rain cloud systems in israel. *J. Appl. Meteorol.,* **35 (9)**, 1502–1510, doi: 10.1175/1520-0450(1996)035<1502:SEIODD>2.0.CO;2.

Rothermel, R. C., 1972: A mathematical model for predicting fire spread in wildland fuels. *Res. Pap. INT-115. Ogden, UT: U.S. Department of Agriculture, Intermountain Forest and Range Experiment Station. 40 p.,* **115**.

Rothermel, R. C., 1983: How to predict the spread and intensity of forest and range fires. Tech. rep., Ogden, UT. doi: 10.2737/INT-GTR-143.

Ruminski, M., S. Kondragunta, R. Draxler, and J. Zeng, 2006: Recent changes to the hazard mapping system. *15th International Emission Inventory Conf,* **(Reinventing Inventories)**.

Sacks, J. D., L. W. Stanek, T. J. Luben, D. O. Johns, B. J. Buckley, J. S. Brown, and M. Ross, 2011: Particulate matter-induced health effects: who is susceptible? *Environ. Health Perspect.*, **119 (4)**, 446–454, doi: 10.1289/ehp.1002255.

Saide, P. E., et al., 2015: Revealing important nocturnal and day-to-day variations in fire smoke emissions through a multiplatform inversion. *Geophys. Res. Lett.*, **42 (9)**, 2015GL063 737, doi: 10.1002/2015GL063737.

Schaaf, M. D., D. V. Sandberg, M. D. Schreuder, and C. L. . Riccardi, 2007: A conceptual framework for ranking crown fire potential in wildland fuelbeds. *Can. J. For. Res.*, **37**, 2464–2478.

Scheller, R., A. Kretchun, T. J. Hawbaker, and P. D. Henne, 2019: A landscape model of variable social-ecological fire regimes. *Ecol. Modell.*, **401**, 85–93, doi: 10.1016/j.ecolmodel.2019.03.022.

Scholze, M., W. Knorr, N. W. Arnell, and I. C. Prentice, 2006: A climate-change risk analysis for world ecosystems. *Proc. Natl. Acad. Sci. U. S. A.*, **103 (35)**, 13 116–13 120.

Schulzweida, U., 2019: CDO user guide. doi: 10.5281/zenodo.2558193.

Short, K. C., 2014: A spatial database of wildfires in the united states, 1992-2011. *Earth System Science Data*, **6 (1)**, 1–27, doi: 10.5194/essd-6-1-2014.

Short, K. C., 2015: Spatial wildfire occurrence data for the united states 1992-2013. *Forest Service Research Data Archive*, **3rd Edition**.

Simms, D. L. and M. Law, 1967: The ignition of wet and dry wood by radiation. *Combust. Flame*, **11 (5)**, 377–388, doi: 10.1016/0010-2180(67)90058-2.

Sing, D. and C. F. Sing, 2010: Impact of direct soil exposures from airborne dust and geophagy on human health. *Int. J. Environ. Res. Public Health*, **7 (3)**, 1205–1223, doi: 10.3390/ijerph7031205.

Spracklen, D. V., L. J. Mickley, J. A. Logan, R. C. Hudman, R. Yevich, M. D. Flannigan, and A. L. Westerling, 2009: Impacts of climate change from 2000 to 2050 on wildfire activity and carbonaceous aerosol concentrations in the western united states. *J. Geophys. Res.*, **114 (D20)**, D20 301, doi: 10.1029/2008JD010966.

Srock, A. F., J. J. Charney, B. E. Potter, and S. L. Goodrick, 2018: The Hot-Dry-Windy index: A new fire weather index. *Atmosphere*, **9 (7)**, 279, doi: 10.3390/atmos9070279.

Steenburgh, W. J., W. James Steenburgh, J. D. Massey, and T. H. Painter, 2012: Episodic dust events of utah's wasatch front and adjoining region. 1654–1669 pp., doi: 10.1175/jamc-d-12-07.1.

Steenland, K. and E. Ward, 2014: Silica: a lung carcinogen. *CA Cancer J. Clin.*, **64 (1)**, 63–69, doi: 10.3322/caac.21214.

Stocks, B. J., T. J. Lynham, B. D. Lawson, M. E. Alexander, C. E. V. Wagner, R. S. McAlpine, and D. E. Dube, 1989: Canadian forest fire danger rating system: an overview. *For. Chron.*, **65 (4)**, 258–265.

Swann, A. L. S., F. M. Hoffman, C. D. Koven, and J. T. Randerson, 2016: Plant responses to increasing CO2 reduce estimates of climate impacts on drought severity. *Proc. Natl. Acad. Sci. U. S. A.*, **113 (36)**, 10 019–10 024, doi: 10.1073/pnas.1604581113.

Syphard, A. D. and J. E. Keeley, 2015: Location, timing and extent of wildfire vary by cause of ignition. *Int. J. Wildland Fire*, **24 (1)**, 37–47, doi: 10.1071/WF14024.

Syphard, A. D., J. E. Keeley, and J. T. Abatzoglou, 2017a: Trends and drivers of fire activity vary across california aridland ecosystems. *J. Arid Environ.*, **144**, 110–122, doi: 10.1016/j.jaridenv.2017.03.017.

Syphard, A. D., J. E. Keeley, A. H. Pfaff, and K. Ferschweiler, 2017b: Human presence diminishes the importance of climate in driving fire activity across the united states. *Proc. Natl. Acad. Sci. U. S. A.*, **114 (52)**, 13 750–13 755, doi: 10.1073/pnas.1713885114.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: A summary of the CMIP5 experiment design. *Bull. Am. Meteorol. Soc.*, **93**, 485–498., doi: 10.1175/BAMS-D-11-00094.1.

Taylor, S. R., 1985: *The Continental Crust: Its Composition and Evolution*. Blackwell Scientific Publications.

Tedim, F., et al., 2018: Defining extreme wildfire events: Difficulties, challenges, and impacts. *Fire*, **1 (1)**, 9, doi: 10.3390/fire1010009.

Thomas, D. S. and D. T. Butry, 2012: Wildland fires within municipal jurisdictions. *J. For.*, **110 (January/February 2012)**.

Tibshirani, R., 1996: Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Series B Stat. Methodol.*, **58 (1)**, 267–288.

Tong, D. Q., J. X. L. Wang, T. E. Gill, H. Lei, and B. Wang, 2017: Intensified dust storm activity and valley fever infection in the southwestern united states. *Geophys. Res. Lett.*, **44 (9)**, 4304–4312, doi: 10.1002/2017GL073524.

Urbanski, S. P., M. C. Reeves, R. E. Corley, R. P. Silverstein, and W. M. Hao, 2018: Contiguous united states wildland fire emission estimates during 2003–2015. doi: 10.5194/essd-10-2241-2018.

USFS, 2015: The rising cost of fire operations: Effects on the forest service's Non-Fire work.

Val Martin, M., C. L. Heald, J.-F. Lamarque, S. Tilmes, L. K. Emmons, and B. A. Schichtel, 2015: How emissions, climate, and land use change will impact mid-century air quality over the united states: a focus on effects at national parks. *Atmos. Chem. Phys.*, **15 (5)**, 2805–2823, doi: 10.5194/acp-15-2805-2015.

van der Sleen, P., et al., 2014: No growth stimulation of tropical trees by 150 years of CO2 fertilization but water-use efficiency increased. *Nat. Geosci.*, **8**, 24, doi: 10.1038/ngeo2313.

van der Werf, G. R., et al., 2010: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmos. Chem. Phys.*, **10 (23)**, 11 707–11 735, doi: 10.5194/acp-10-11707-2010.

Van Donkelaar, A., R. V. Martin, M. Brauer, R. Kahn, R. Levy, C. Verduzco, and P. J. Villeneuve, 2010: Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ. Health Perspect.*, **118 (6)**, 847–855.

van Donkelaar, A., R. V. Martin, A. N. Pasch, J. J. Szykman, L. Zhang, Y. X. Wang, and D. Chen, 2012: Improving the accuracy of daily satellite-derived ground-level fine aerosol concentration estimates for north america. *Environ. Sci. Technol.*, **46 (21)**, 11 971–11 978, doi: 10.1021/es3025319.

Vedal, S. and S. J. Dutton, 2006: Wildfire air pollution and daily mortality in a large urban area. *Environ. Res.*, **102 (1)**, 29–35, doi: 10.1016/j.envres.2006.03.008.

Wagner, C. E. V., 1977: Conditions for the start and spread of crown fire. *Can. J. For. Res.*, **7 (1)**, 23–34, doi: 10.1139/x77-004.

Wagner, C. E. V., 1998: Modelling logic and the canadian forest fire behavior prediction system. *For. Chron.*, **74 (1)**, 50–52, doi: 10.5558/tfc74050-1.

Wang, R., et al., 2015: Sources, transport and deposition of iron in the global atmosphere. *Atmos. Chem. Phys.*, **15 (11)**, 6247–6270, doi: 10.5194/acp-15-6247-2015.

Ward, T. J., R. F. Hamilton, R. W. Dixon, M. Paulsen, and C. D. Simpson, 2006: Characterization and evaluation of smoke tracers in PM: Results from the 2003 montana wildfire season. *Atmos. Environ.*, **40 (36)**, 7005–7017, doi: 10.1016/j.atmosenv.2006.06.034.

Warren, J. M., E. Pötzelsberger, S. D. Wullschleger, P. E. Thornton, H. Hasenauer, and R. J. Norby, 2011: Ecohydrologic impact of reduced stomatal conductance in forests exposed to elevated CO2. *Ecohydrol.*, **4 (2)**, 196–210, doi: 10.1002/eco.173.

Warren, S. G. and R. E. Brandt, 2008: Optical constants of ice from the ultraviolet to the microwave: A revised compilation. *J. Geophys. Res.*, **113 (D14)**, D13 203, doi: 10.1029/2007JD009744.

Wells, K. C., M. Witek, P. Flatau, S. M. Kreidenweis, and D. L. Westphal, 2007: An analysis of seasonal surface dust aerosol concentrations in the western US (2001–2004): Observations and model predictions. *Atmos. Environ.*, **41 (31)**, 6585–6597, doi: 10.1016/j.atmosenv.2007.04.034.

Westerling, A. L., 2016: Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **371 (1696)**, 20150 178, doi: 10.1098/rstb.2015.0178.

Westerling, A. L., T. Brown, T. Schoennagel, T. Swetnam, M. Turner, and T. Veblen, 2014: *Briefing: Climate and wildfire in western U.S. forests*, 81–102.

Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam, 2006: Warming and earlier spring increase western U.S. forest wildfire activity. *Science*, **313 (5789)**, 940–943, doi: 10.1126/science.1128834.

Westerling, L., M. G. Turner, E. A. H. Smithwick, W. H. Romme, and M. G. Ryan, 2011: Continued warming could transform greater yellowstone fire regimes by mid-21st century. *Proc. Natl. Acad. Sci. U. S. A.*, **108 (32)**, 13 165–13 170, doi: 10.1073/pnas.1110199108.

WFAS, ????: Fire danger rating. Accessed: 2018-8-30, `https://www.wfas.net/index.php/fire-danger-rating-fire-potential--danger-32`.

Wiedinmyer, C., S. K. Akagi, R. J. Yokelson, L. K. Emmons, J. A. Al-Saadi, J. J. Orlando, and A. J. Soja, 2011: The fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development*, **4 (3)**, 625–641, doi: 10.5194/gmd-4-625-2011.

Wiedinmyer, C., B. Quayle, C. Geron, A. Belote, D. McKenzie, X. Zhang, S. O'Neill, and K. K. Wynne, 2006: Estimating emissions from fires in north america for air quality modeling. *Atmos. Environ.*, **40 (19)**, 3419–3432, doi: 10.1016/j.atmosenv.2006.02.010.

Wilson, G., et al., 2014: Best practices for scientific computing. *PLoS Biol.*, **12 (1)**, e1001 745, doi: 10.1371/journal.pbio.1001745.

Woodward, S., D. L. Roberts, and R. A. Betts, 2005: A simulation of the effect of climate change-induced desertification on mineral dust aerosol: EFFECTS OF DESERTIFICATION ON DUST. *Geophys. Res. Lett.*, **32 (18)**, doi: 10.1029/2005GL023482.

Wu, J., A. M. Winer, and R. J. Delfino, 2006: Exposure assessment of particulate matter air pollution before, during, and after the 2003 southern california wildfires. *Atmos. Environ.*, **40 (18)**, 3333–3348.

Yoon, J.-H., B. Kravitz, P. J. Rasch, S.-Y. Simon Wang, R. R. Gillies, and L. Hipps, 2015: Extreme fire season in california: A glimpse into the future? *Bull. Am. Meteorol. Soc.*, **96 (12)**, S5–S9, doi: 10.1175/BAMS-D-15-00114.1.

Yue, X., L. J. Mickley, and J. A. Logan, 2014: Projection of wildfire activity in southern california in the mid-twenty-first century. *Clim. Dyn.*, **43 (7-8)**, 1973–1991, doi: 10.1007/s00382-013-2022-3.

Yue, X., L. J. Mickley, J. A. Logan, and J. O. Kaplan, 2013a: Ensemble projections of wildfire activity and carbonaceous aerosol concentrations over the western united states in the mid-21st century. *Atmos. Environ.*, **77**, 767–780, doi: 10.1016/j.atmosenv.2013.06.003.

Yue, X., L. J. Mickley, J. A. Logan, and J. O. Kaplan, 2013b: Ensemble projections of wildfire activity and carbonaceous aerosol concentrations over the western united states in the mid-21st century. *Atmos. Environ.*, **77**, 767–780, doi: 10.1016/j.atmosenv.2013.06.003.

Yue, X. and N. Unger, 2018: Fire air pollution reduces global terrestrial productivity. *Nat. Commun.*, **9 (1)**, 5413, doi: 10.1038/s41467-018-07921-4.

Zhao, M. and S. W. Running, 2010: Drought-Induced reduction in global terrestrial net primary production from 2000 through 2009. *Science*, **329 (5994)**, 940–943, doi: 10.1126/science.1192666.
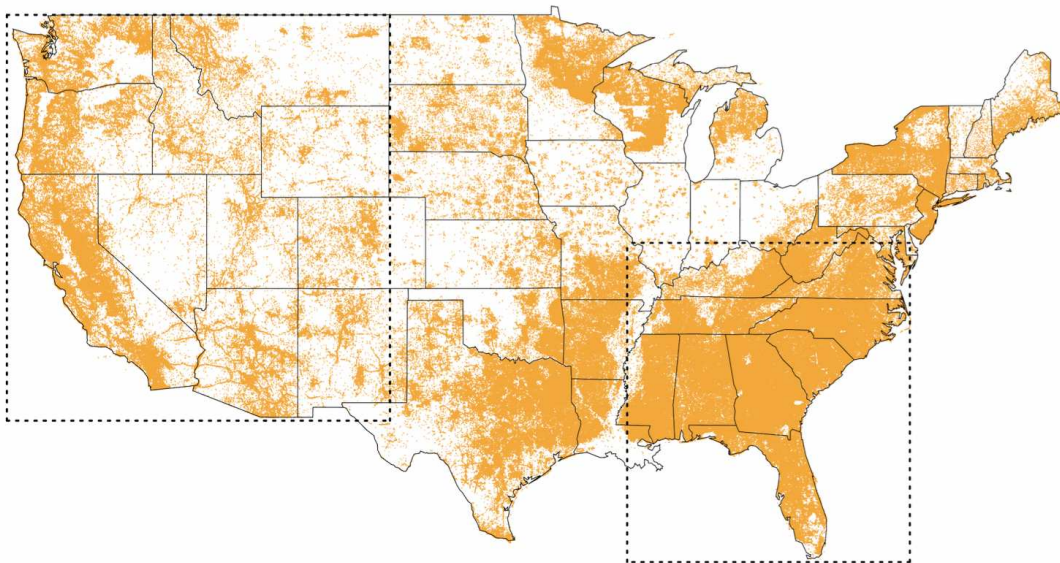
A1 SUPPLEMENTAL INFORMATION FOR CHAPTER 2
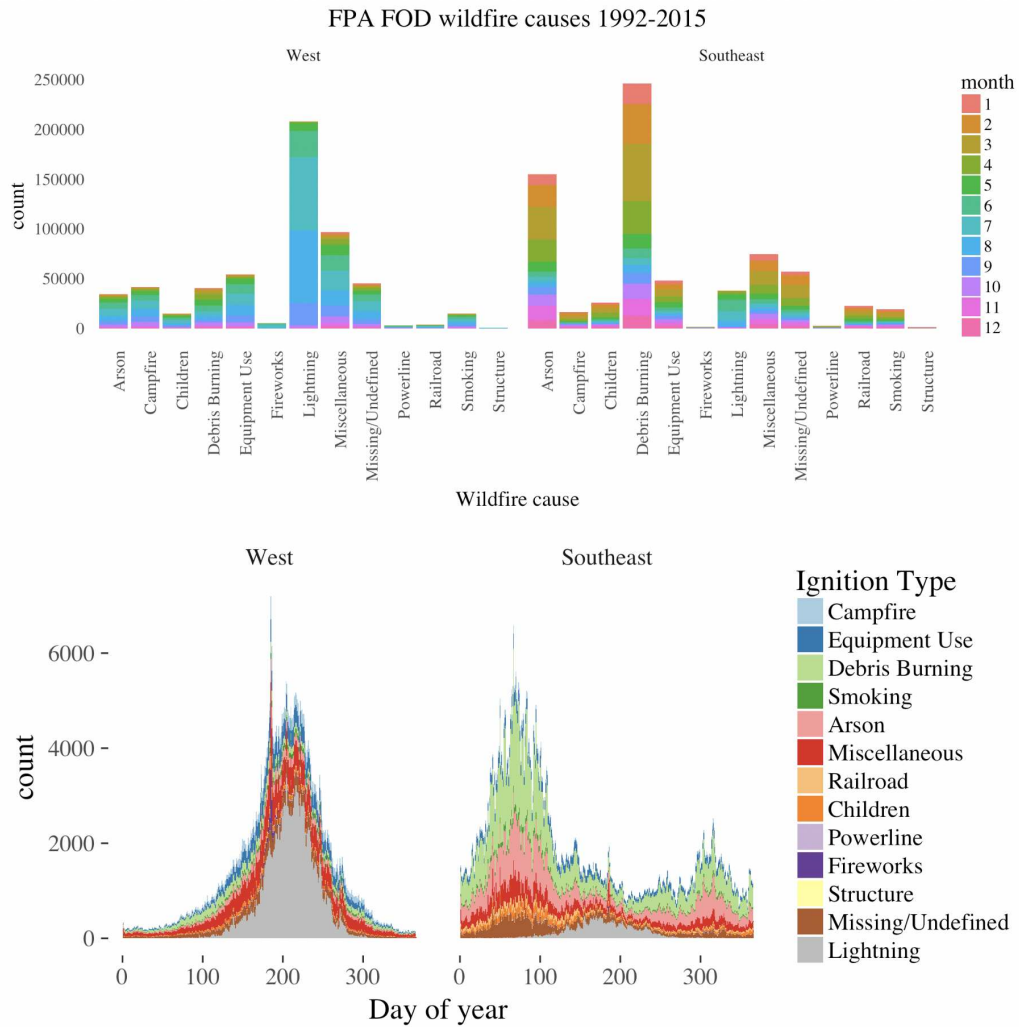
FIG. A1. (Top) Shows the count for wildfire ignition types in the west and southeast, the bars are color coded by the wildfire discovery month. (Bottom) Shows the count for each ignition type vs. the Day of the year it was discovered in the west and southeast region between 1992 and 2015.
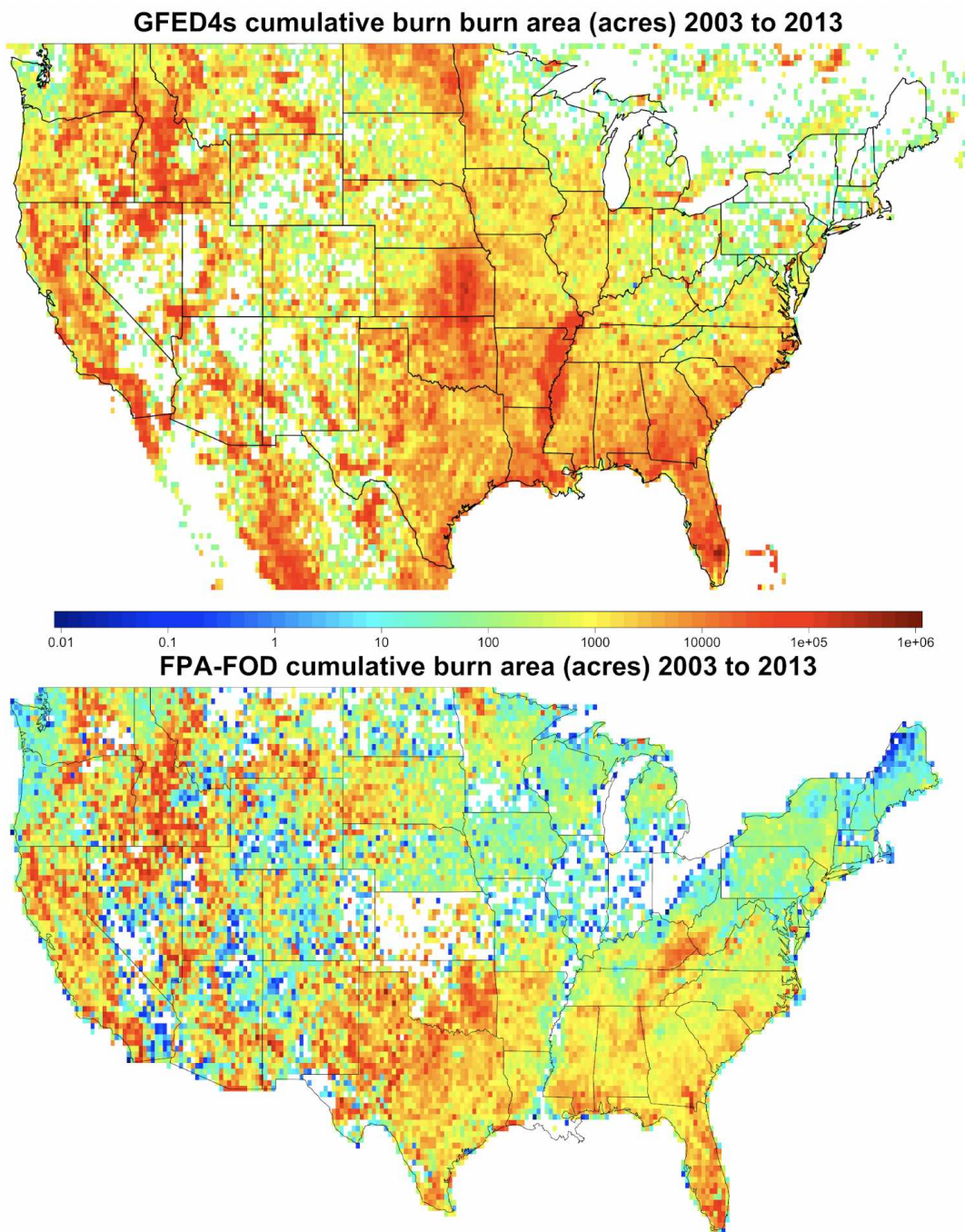
Lightning-ignited Wildfires | 1992 - 2015
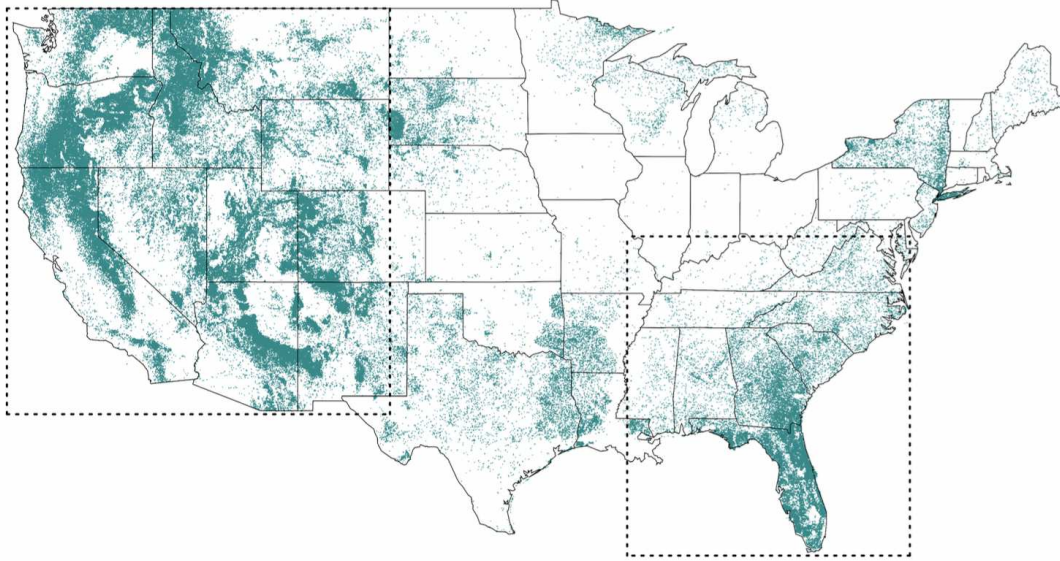
Human-ignited Wildfires | 1992 - 2015

FIG. A2. (Top) Shows the count for wildfire ignition types in the west and southeast, the bars are color coded by the wildfire discovery month. (Bottom) Shows the count for each ignition type vs. the Day of the year it was discovered in the west and southeast region between 1992 and 2015.

FIG. A3. Total burn area (m2) for GFEDv4s (top) and FPA FOD (bottom) for all months between 2003 and 2013 on a 0.25° x 025° grid. GFEDv4s is available 2003 onwards, FPA FOD is available from 1992-2015, however, these figures were made before 2014 and 2015 came available. We expect the patterns to remain similar for 2014 and 2015 data. This figure demonstrates significant differences in burn area estimates between an all lands satellite-based burn area product (GFEDv4s) and the FPA FOD, which only accounts for wildfires.
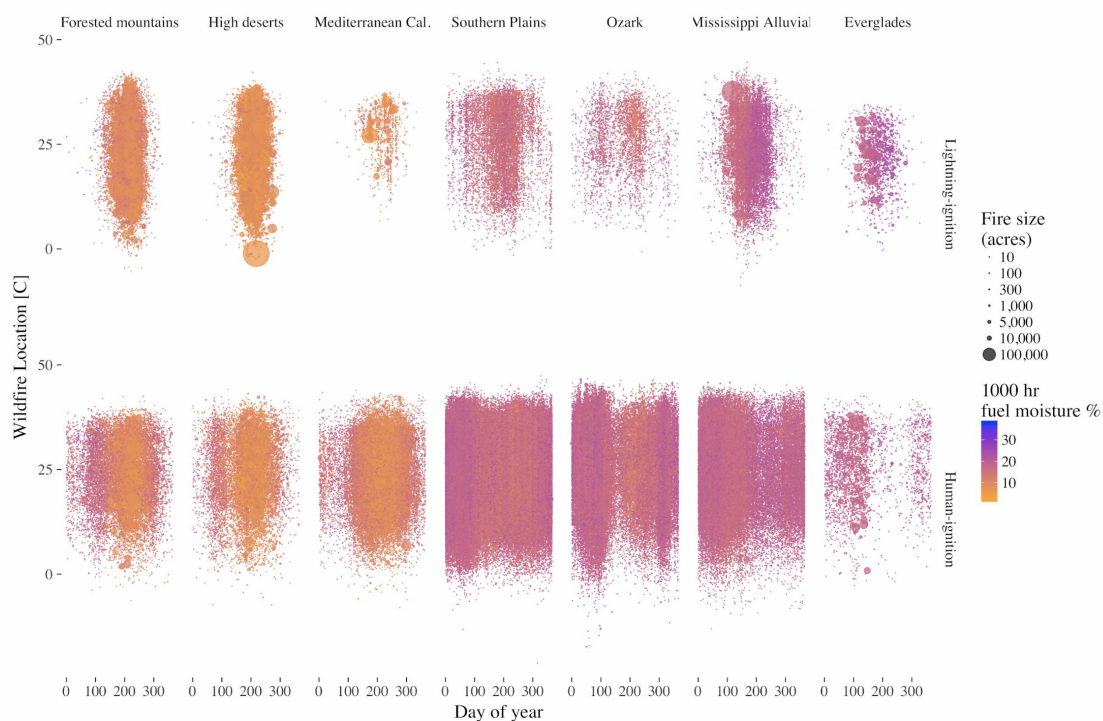
FIG. A4. This figure shows the size (dot area proportional to acres burned) and 1000-hour dead fuel moisture (dot color) of FPA FOD wildfires plotted on a temperature vs. day of year axis. The top row shows lightning-ignited wildfires, the bottom human-ignited, the columns show the seven west and southeast ecoregions of interest. For visual clarity, wildfires smaller than 1 acre are excluded from this figure. The fire size legend reflects the sizes of different wildfire size classes defined by the National Wildfire Coordination Group (NWCG). For regular non-leap-day years, Julian Day 100 is April 10th, 150 is May 30th, 200 is July 19, 250 is September 7th, and 300 is Oct 27th.
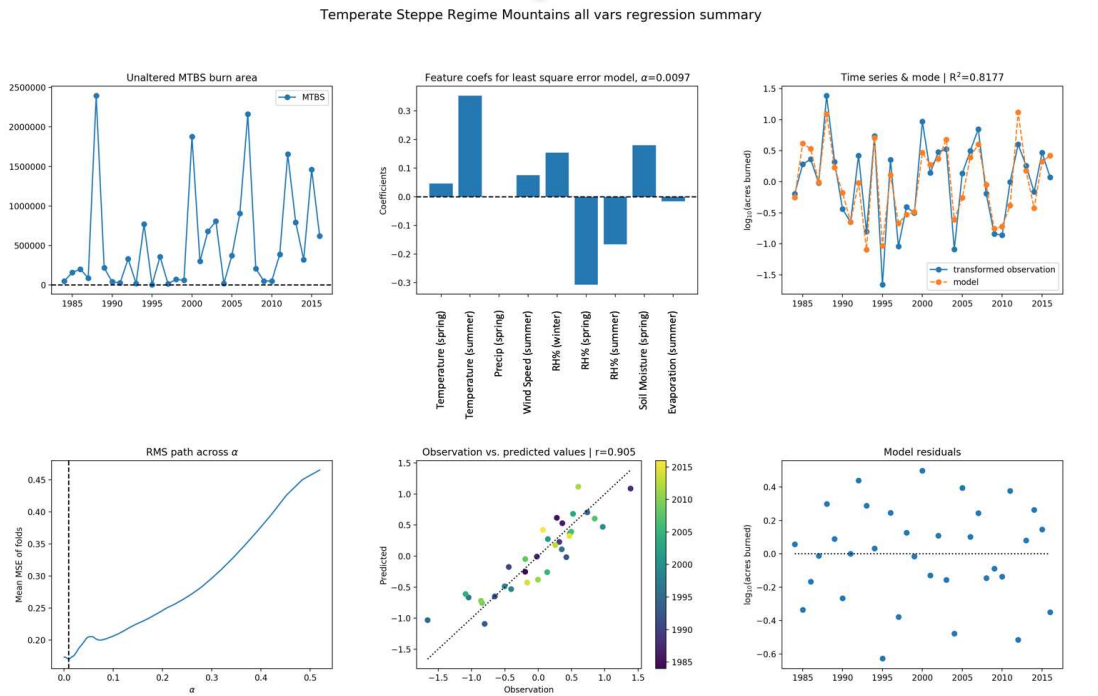
FIG. A5. Summary of the Lasso-All (access to all candidate variables) relaxed-Lasso regression for the Temperate Steppe Regime Mountains. The regression shown is for the top 50% of candidate variables. Top left: Monitoring Trends in Burn Severity (MTBS) June through August acres burned time series for the Temperate Steppe Regime Mountains. Top Middle: The coefficients of the relaxed Lasso. Top right: The time series used for regression (blue) and the regression-model estimate (orange). Bottom left: Mean square error across folds (years left out) for each value of alpha (horizontal axis) tested. The value of alpha with the lowest MSE is what was selected for the final regression model, this value is indicated with a black vertical dashed line. Bottom middle: Transformed observed burn area values vs. regression estimated value (associated with curves in top right). Bottom right: relaxed-Lasso regression residuals.

Temperate Steppe Regime Mountains atmos-centric regression summary
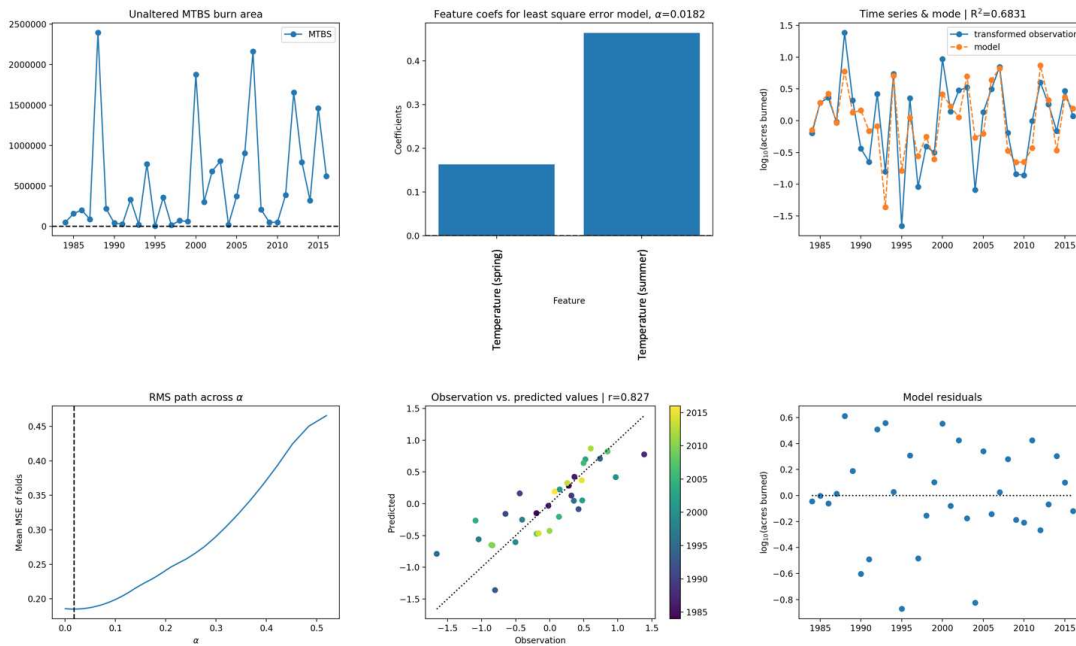
FIG. A6. The same as Figure A5, except this summary is for the Lasso-Temperature (access to only temperature) relaxed-Lasso regression for the Temperate Steppe Regime Mountains.
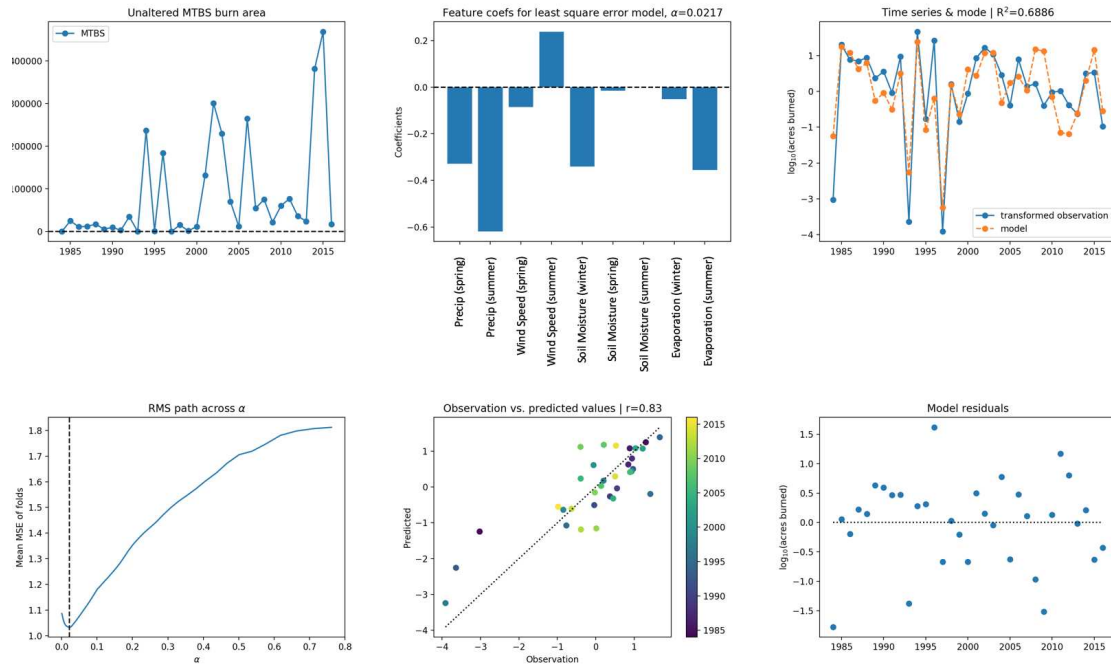
FIG. A7. Summary of the Lasso-All (access to all candidate variables) relaxed-Lasso regression for the Marine Regime Mountains. The regression shown is for the top 50% of candidate variables. Top left: Monitoring Trends in Burn Severity (MTBS) June through August acres burned time series for the Marine Regime Mountains. Top Middle: The coefficients of the relaxed Lasso. Top right: The time series used for regression (blue) and the regression-model estimate (orange). Bottom left: Mean square error across folds (years left out) for each value of alpha (horizontal axis) tested. The value of alpha with the lowest MSE is what was selected for the final regression model, this value is indicated with a black vertical dashed line. Bottom middle: Transformed observed burn area values vs. regression estimated value (associated with curves in top right). Bottom right: relaxed-Lasso regression residuals.
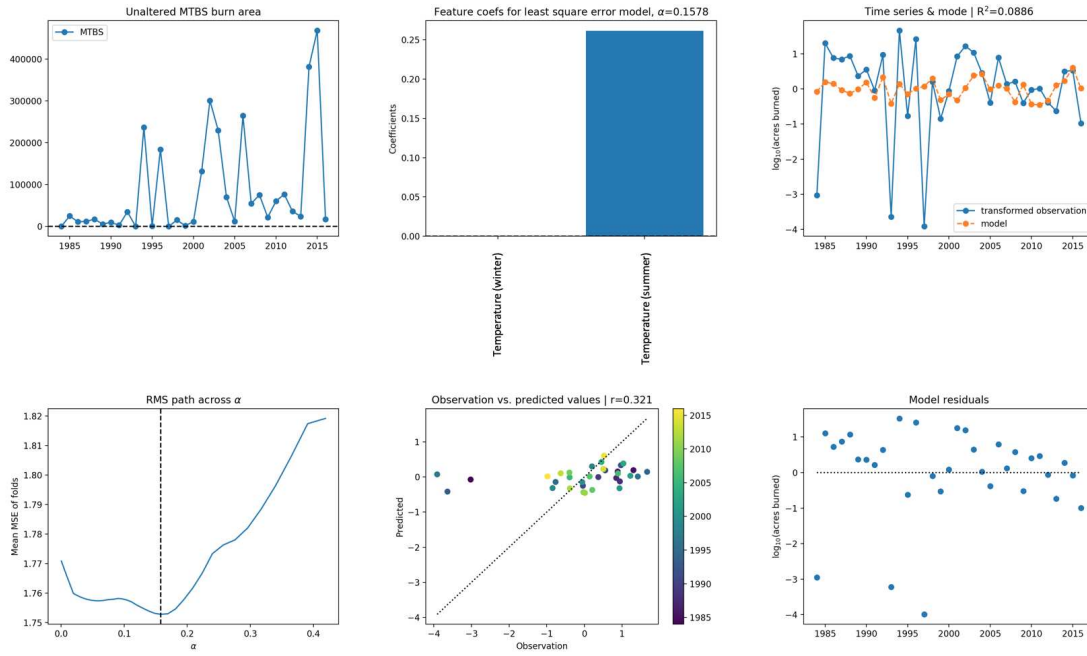
FIG. A8. The same as Figure A7, except this summary is for the Lasso-Temperature (access to only temperature) relaxed-Lasso regression for the Marine Regime Mountains.

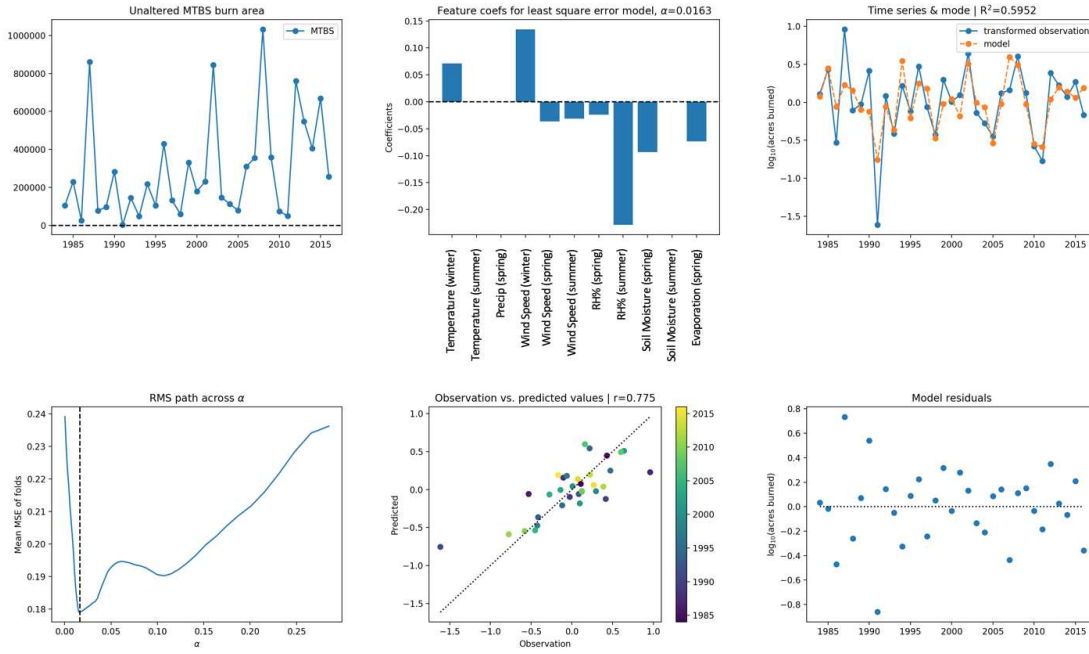Mediterranean Regime Mountains all vars regression summary

FIG. A9. Summary of the Lasso-All (access to all candidate variables) relaxed-Lasso regression for the Mediterranean Regime Mountains. The regression shown is for the top 50% of candidate variables. Top left: Monitoring Trends in Burn Severity (MTBS) June through August acres burned time series for the Mediterranean Regime Mountains. Top Middle: The coefficients of the relaxed Lasso. Top right: The time series used for regression (blue) and the regression-model estimate (orange). Bottom left: Mean square error across folds (years left out) for each value of alpha (horizontal axis) tested. The value of alpha with the lowest MSE is what was selected for the final regression model, this value is indicated with a black vertical dashed line. Bottom middle: Transformed observed burn area values vs. regression estimated value (associated with curves in top right). Bottom right: relaxed-Lasso regression residuals.
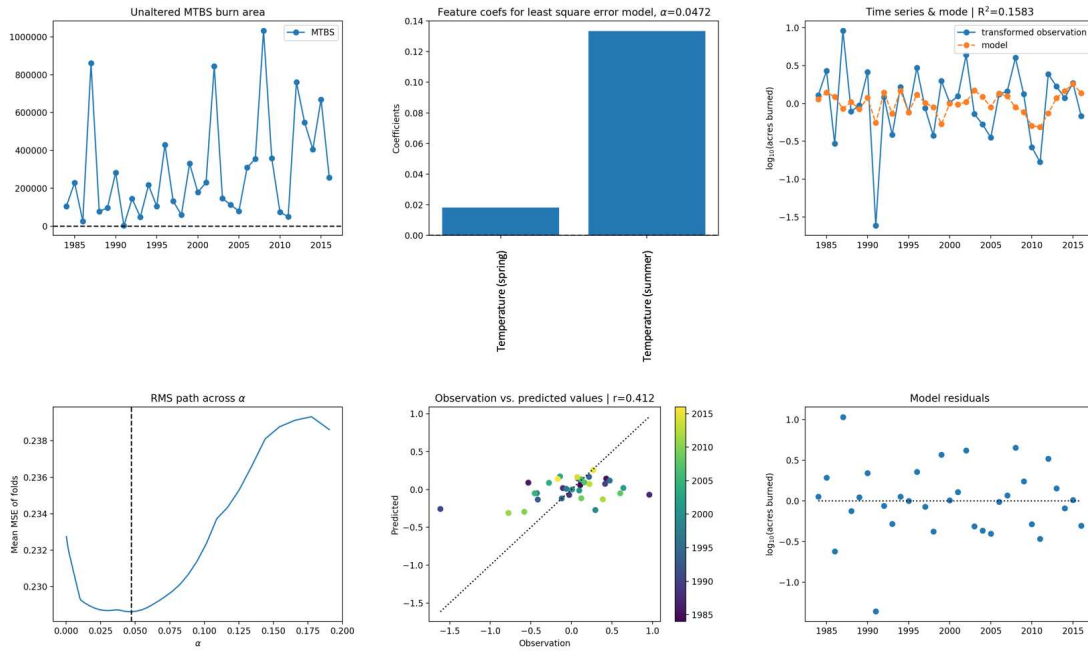
FIG. A10. The same as Figure A9, except this summary is for the Lasso-Temperature (access to only temperature) relaxed-Lasso regression for the Mediterranean Regime Mountains.
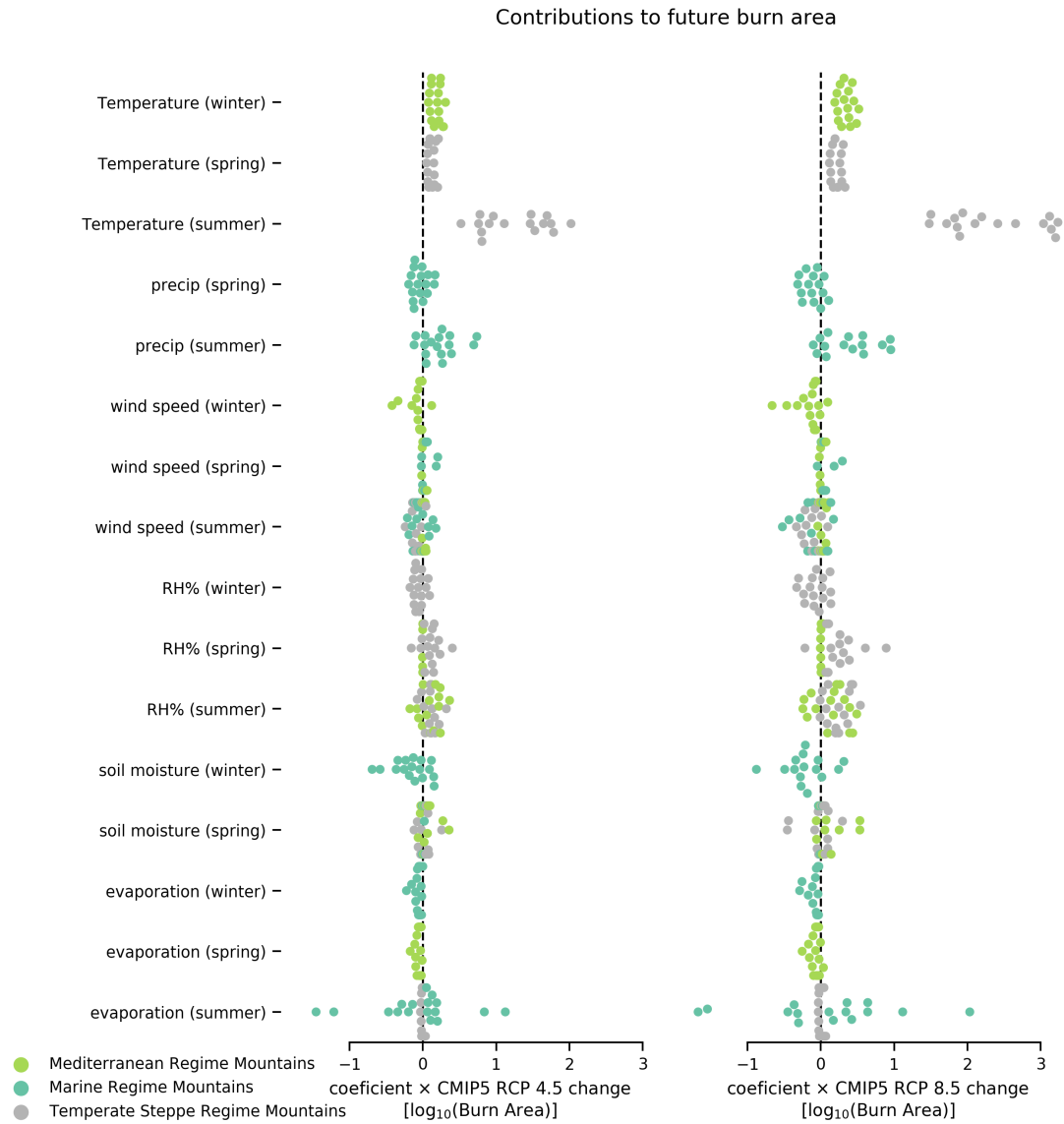
FIG. A11. Change in burn area estimated using regression coefficient and model estimated change in that value by 2070-2099 for RCP 4.5 and 8.5. The horizontal axis shows the expected change in burn area, each dot is a single models estimate, based on the amount a variable changes from the historical period.

Site	Code	Dataset	State	County	Latitude	Longitude	Elevation	StartDate	EndDate	NumPOCs
Bandelier NM	BAND1	IMPFSPED	NM	35028	35.7797	-106.26640	1988	03/02/1988	06/28/2018	1
Bryce Canyon NP	BRCA1	IMPFSPED	UT	49017	37.6184	-112.17360	2481	03/02/1988	06/28/2018	1
Canyonlands NP	CANY1	IMPFSPED	UT	49037	38.4587	-109.82100	1798	03/02/1988	06/28/2018	1
Chiricahua NM	CHIR1	IMPFSPED	AZ	4003	32.0094	-109.38900	1554	03/02/1988	06/28/2018	1
Gila Wilderness	GICL1	IMPFSPED	NM	35003	33.2204	-108.23510	1775	03/02/1994	06/28/2018	1
Great Basin NP	GRBA1	IMPFSPED	NV	32033	39.0052	-114.21610	2065	03/04/1992	06/28/2018	1
Great Sand Dunes NM	GRSA1	IMPFSPED	CO	8003	37.7249	-105.51850	2498	03/02/1988	06/28/2018	1
Guadalupe Mountains NP	GUMO1	IMPFSPED	TX	48109	31.8330	-104.80940	1672	03/02/1988	06/28/2018	1
Meadview	MEAD1	IMPFSPED	AZ	4015	36.0193	-114.06840	902	09/04/1991	06/28/2018	1
Mesa Verde NP	MEVE1	IMPFSPED	CO	8083	37.1984	-108.49070	2172	03/02/1988	06/28/2018	2
Mount Zirkel Wilderness	MOZI1	IMPFSPED	CO	8057	40.5383	-106.67660	3243	06/01/1994	06/28/2018	1
Petrified Forest NP	PEFO1	IMPFSPED	AZ	4017	35.0777	-109.76920	1766	03/02/1988	06/28/2018	1
Rocky Mountain NP	ROMO1	IMPFSPED	CO	8069	40.2783	-105.54570	2760	09/01/1990	06/28/2018	1
Saguaro NM	SAGU1	IMPFSPED	AZ	4019	32.1746	-110.73710	941	06/01/1988	06/28/2018	1
Tonto NM	TONT1	IMPFSPED	AZ	4007	33.6548	-111.10680	775	03/02/1988	06/28/2018	1
Weminuche Wilderness	WEMI1	IMPFSPED	CO	8111	37.6594	-107.79988	2750	03/02/1988	06/28/2018	1
White River NF	WHRI1	IMPFSPED	CO	8097	39.1536	-106.82090	3413	06/02/1993	06/28/2018	1

FIG. A12. Summary of IMPROVE sites used in this analysis. Saguaro NM (sitecode SAGU1), missing measurements between 1995 and 1999 and Meadview Arizona (sitecode MEAD1), missing data between 1995 and 2003.
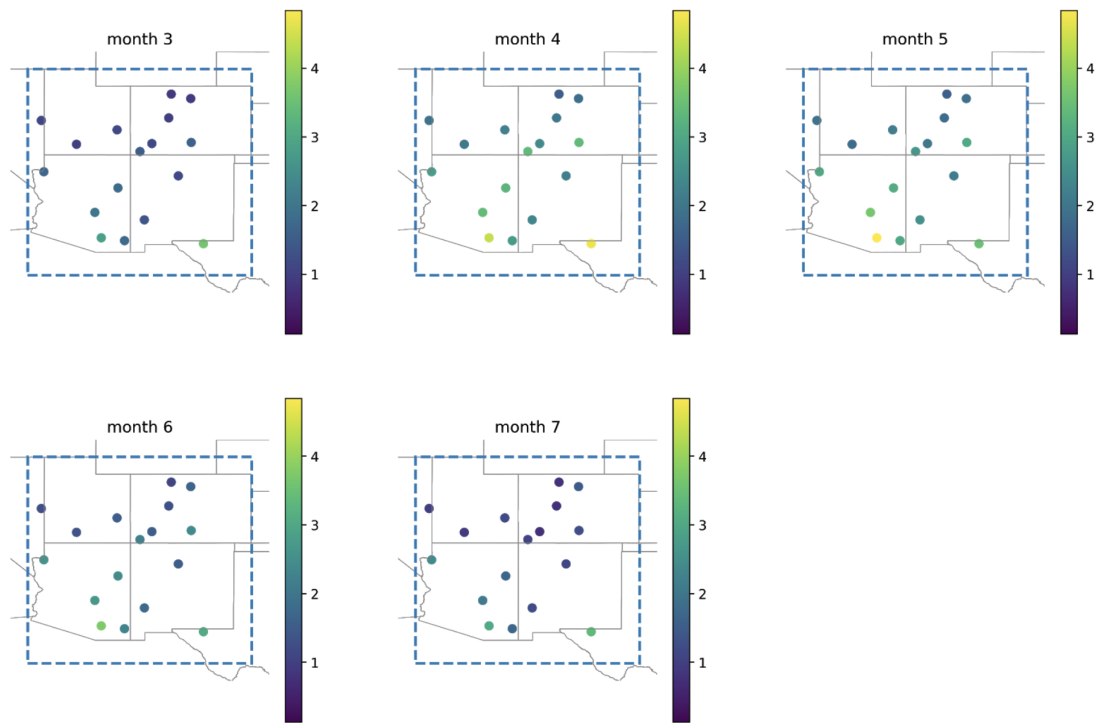
FIG. A13. Monthly mean dust concentration (μg m^{-3}) values values for monitors used in this analysis. The plotted monthly means for each monitor was calculated using all data available between 1995 and 2015. Monitors further south tend to measure higher dust concentrations.
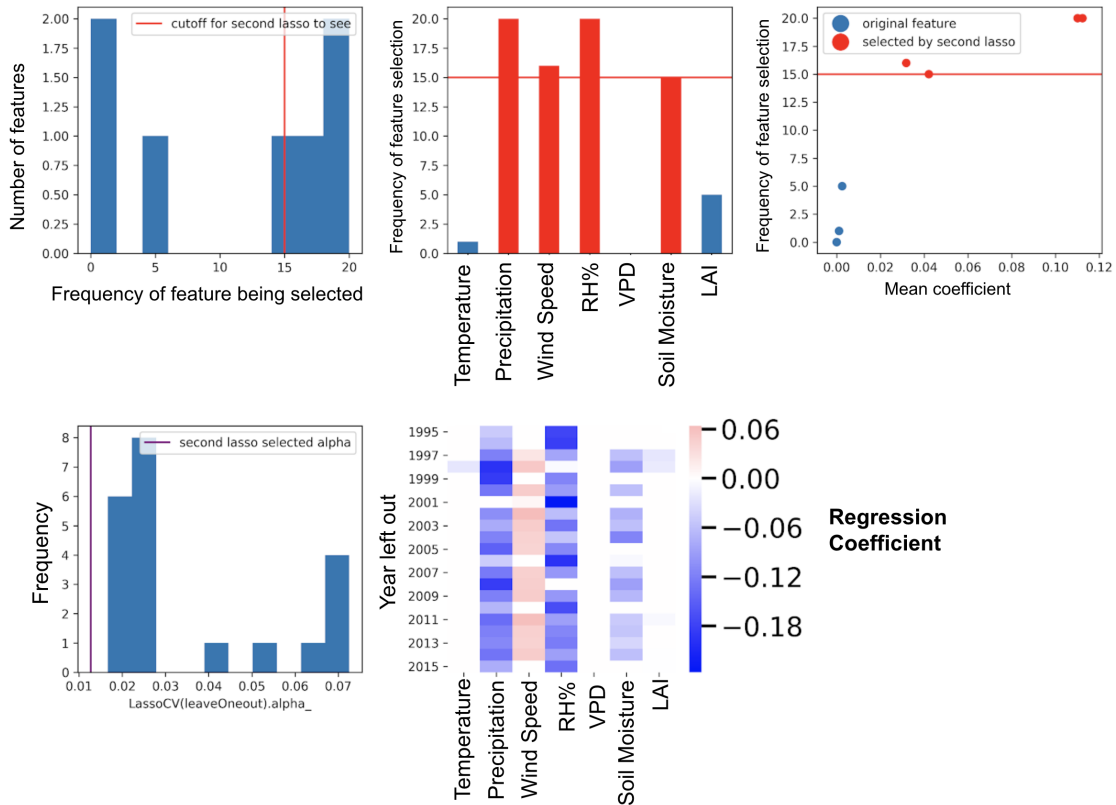
FIG. A14. Figure S3: Overview of Relaxed-Lasso variable selection output. Topleft: The distribution of the number of times a variable was selected in the year loop (variable selection part of relaxo). Top middle: The number of times candidate variables were selected in the variable selection loop. The horizontal red lines shows the threshold that must be exceeded in order for the variable to be selected for the final Lasso fit. Topright: The relationship between the number of times a variable was selected against the mean coefficient value for selected variables. Bottom-left: The blue bars show the distribution of the α values for the external variable selection loop, the vertical purple line shows the α value for the lasso fit after the variable selection loop. Bottom middle: The regression coefficient value for the variables for the years left out variable selection loop.
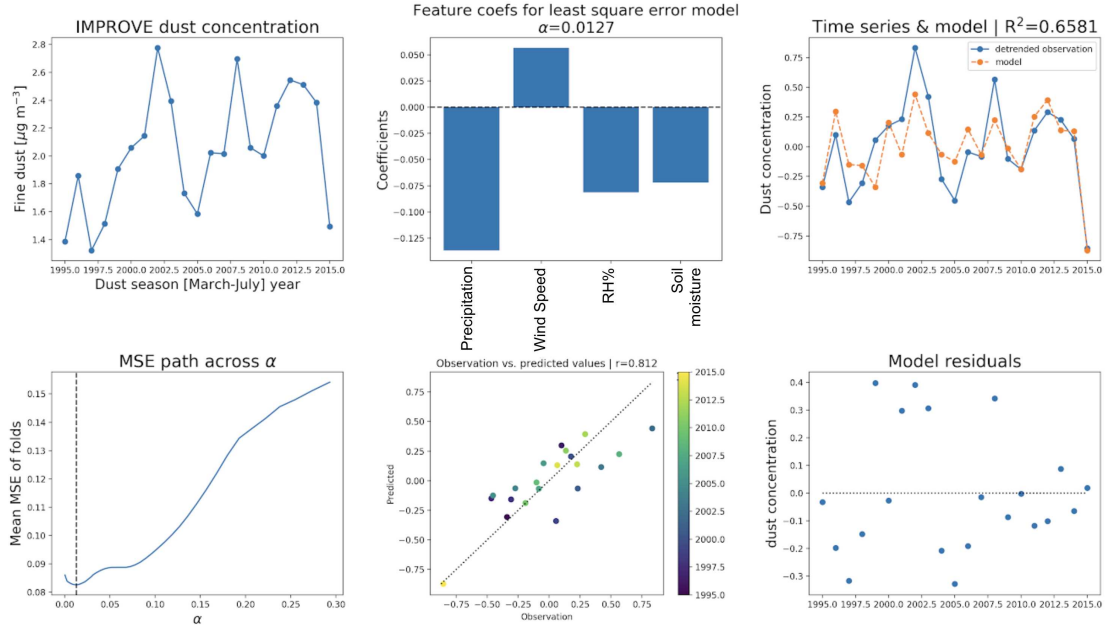
FIG. A15. Overview of the final step of the Relaxed-Lasso regression, performing Lasso regression using the variables selected in the variable selection step (shown in Figure AA14). Top Left: Southwest dust season (March through July) concentration time series. Top Middle: The regression coefficients for the fit Relaxed-Lasso. Top Right: The zero-centered detrended target function (dust concentration) and regression fit. The R^2 value for the fit is 0.6581. Bottom Left: The mean square error (MSE) for a given value of α. The dashed vertical line shows the final value of α selected for the regression. Bottom Middle: The value predicted by the regression vs. the target function. Bottom Right: The residuals for the regression.